

This page Is Inserted by IFW Operations
And is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of
The original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
Please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
17 October 2002 (17.10.2002)

PCT

(10) International Publication Number
WO 02/081731 A2

(51) International Patent Classification⁷: **C12Q**
(21) International Application Number: PCT/US02/01222
(22) International Filing Date: 29 January 2002 (29.01.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/774,528 30 January 2001 (30.01.2001) US

(71) Applicant (for all designated States except US): **HYSEQ, INC.** [US/US]; 670 Almanor Avenue, Sunnyvale, CA 94086 (US).

(71) Applicant (for US only): **GOODRICH, Ryle, W.** [US/US]; 4896 Sandy Lane, San Jose, CA 95124 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **TANG, Tom, Y.** [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **LIU, Chenghua** [CN/US]; 1125 Ranchero Way #14, San Jose, CA 95117 (US). **ZHOU, Ping** [CN/US]; 7595 New-castle Drive, Cupertino, CA 95014 (US). **ASUNDI, Vinod** [US/US]; 709 Foster City Boulevard, Foster City, CA 94404 (US). **ZHANG, Jie** [CN/US]; 4930 Poplar Terrace, Campbell, CA 95008 (US). **ZHAO, Qing, A.** [CN/US]; 1556 Kooser Road, San Jose, CA 95118 (US). **REN, Feiyan** [US/US]; 7703 Oak Meadow Court, Cupertino, CA 95014 (US). **XUE, Aidong, J.** [CN/US]; 1621 S. Mary Avenue, Sunnyvale, CA 94087 (US). **YANG, Yonghong** [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **WEHRMAN, Tom** [US/US]; 3210 CCSR Mol Pharm, 269 W. Campus Drive, Stanford, CA 94305 (US). **WANG,**

Jian-Rui [CN/US]; 744 Stendhal Lane, Cupertino, CA 95014 (US). **WANG, Dunrui** [CN/US]; 12252 Pepper Tree Lane, Poway, CA 92064 (US). **DRMANAC, Radoje, T.** [US/US]; 850 Greenwich Place, Palo Alto, CA 94303 (US).

(74) Agents: **HSI, Petrina, S. et al.**; Hyseq, Inc., 670 Almanor Avenue, Sunnyvale, CA 94085 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- without international search report and to be republished upon receipt of that report
- with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

WO 02/081731 A2

NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by
5 such polynucleotides, along with uses for these polynucleotides and proteins, for example
in therapeutic, diagnostic and research methods.

2. BACKGROUND

Technology aimed at the discovery of protein factors (including e.g., cytokines,
10 such as lymphokines, interferons, CSFs, chemokines, and interleukins) has matured
rapidly over the past decade. The now routine hybridization cloning and expression
cloning techniques clone novel polynucleotides "directly" in the sense that they rely on
information directly related to the discovered protein (i.e., partial DNA/amino acid
sequence of the protein in the case of hybridization cloning; activity of the protein in the
15 case of expression cloning). More recent "indirect" cloning techniques such as signal
sequence cloning, which isolates DNA sequences based on the presence of a now
well-recognized secretory leader sequence motif, as well as various PCR-based or low
stringency hybridization-based cloning techniques, have advanced the state of the art by
making available large numbers of DNA/amino acid sequences for proteins that are
20 known to have biological activity, for example, by virtue of their secreted nature in the
case of leader sequence cloning, by virtue of their cell or tissue source in the case of
PCR-based techniques, or by virtue of structural similarity to other genes of known
biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications
25 in, for example, diagnostics, forensics, gene mapping; identification of mutations
responsible for genetic disorders or other traits, to assess biodiversity, and to produce
many other types of data and products dependent on DNA and amino acid sequences.

3. SUMMARY OF THE INVENTION

30 The compositions of the present invention include novel isolated polypeptides, novel
isolated polynucleotides encoding such polypeptides, including recombinant DNA

molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

5 The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

10 The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO:
15 1 – 438 and are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, * corresponds to the stop codon.

20 The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO: 1 – 438 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO: 1 – 438. A polynucleotide comprising a nucleotide sequence having at least
25 90% identity to an identifying sequence of SEQ ID NO: 1 – 438 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

30 The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO: 1 – 438. The sequence information can be a segment of any one of SEQ ID NO: 1 – 438 that uniquely identifies or represents the sequence information of SEQ ID NO: 1 – 438.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information is provided on a nucleic acid array to detect the polynucleotide that
5 contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid
10 sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing full-length genes, use for
15 chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO: 1-438 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the
20 nucleic acid sequences of SEQ ID NO: 1-438 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The isolated polynucleotides of the invention include, but are not limited to, a
25 polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO: 1-438; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO: 1-438; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1-438. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under
30 stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO: 1-438; (b) a nucleotide sequence encoding any one of

the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO: 1-438; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein,

and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, *e.g.*, *in situ*

5 hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

10 The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the
15 invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

20 In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for
25 example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form
30 the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The

invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (*e.g.*, bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and detecting the complex by detecting the reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provides methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases or disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Table 2); for which they have a signature region (as set forth in Table 3); or for which they have homology to a gene family (as set forth in Table 4). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

4. DETAILED DESCRIPTION OF THE INVENTION

4.1 DEFINITIONS

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived. The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonucleotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid

which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of
5 nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 9 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50
10 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or
15 amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NOs:1-438.

Probes may, for example, be used to determine whether specific mRNA
20 molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A
25 Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NOs: 1-438. The sequence
30 information can be a segment of any one of SEQ ID NOs: 1-438 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO: 1-438. One such

segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4^{20} possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match ($1+4^{25}$) times the increased probability for mismatch at each nucleotide position (3×25). The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements e.g. repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids, more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol)

and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

The term "variant" (or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, *e.g.*, recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions,

deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, *e.g.*, polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (*e.g.*, nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (*e.g.*, microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (*e.g.*, yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, *e.g.*, *E. coli*, will be free of

glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An
5 expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use in yeast or eukaryotic expression
10 systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

15 The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed.
20 This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be
25 prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (*e.g.*, soluble proteins) or partially (*e.g.*, receptors) from the cell
30 in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted"

proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134-143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

5 Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

10 The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (i.e., hybridization to filter-bound DNA in 0.5 M NaHPO₄, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (i.e., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

15 In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

20 As used herein, "substantially equivalent" or "substantially similar" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (i.e., the number of individual residue
25 substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, e.g., mutant, sequence of the invention varies from a listed
30 sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of

this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more than 5% (95% sequence identity). Substantially equivalent, *e.g.*, mutant, amino acid sequences according to the invention preferably have

5 at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% sequence identity, more preferably at least 98% sequence identity, and most preferably at least 99% sequence identity. Substantially equivalent nucleotide sequence of the invention can have lower percent sequence identities, taking into account, for

10 example, the redundancy or degeneracy of the genetic code. Preferably, the nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least about 95% sequence identity, more preferably at least 98% sequence identity, and

15 most preferably at least 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (*e.g.*, via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity

20 may be determined, *e.g.*, using the Jotun Hein method (Hein, J. (1990) *Methods Enzymol.* 183:626-645). Identity between sequences can also be determined by other methods known in the art, *e.g.* by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

25 The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a

30 virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

4.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO: 1 – 438; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO: 1 – 438; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polynucleotides of any one of SEQ ID NO: 1 - 438. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO: 1 – 438; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 1- 438. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

5 The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic
10 materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO: 1 - 438 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID
15 NO: 1 - 438 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO: 1 - 438 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

 The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpr, and UniGene. The EST sequences can provide identifying
20 sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

 The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, e.g., at least about 65%, at
25 least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99% sequence identity to a polynucleotide recited above.

 Included within the scope of the nucleic acid sequences of the invention are
30 nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO: 1 - 438, or complements thereof, which fragment is

greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that are selective for (i.e. specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided in SEQ ID NO: 1 - 438, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NOs: 1 - 438 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NOs: 1 - 438, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altschul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, *e.g.*, by substituting first with conservative choices (*e.g.*, hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (*e.g.*, hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., *DNA* 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, *Nucleic Acids Res.* 10:6487-6500 (1982). PCR may also be used to

create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ
5 from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., *Gene* 34:315 (1985); and other mutagenesis
10 techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and *Current Protocols in Molecular Biology*, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression of these novel nucleic acids. Such DNA
15 sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

20 The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate
25 polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO: 1-438, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in
30 appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an
5 assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell.

- 10 Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a
15 nucleic acid having any of the nucleotide sequences of SEQ ID NOs: 1 - 438 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NOs: 1 - 438 or a fragment thereof is inserted, in a forward or reverse orientation. In the
20 case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of
25 example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

- 30 The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein

recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, *e.g.*, the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, *e.g.*, stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus*

subtilis, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (*e.g.*, temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intra-muscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

4.3 ANTISENSE

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1 - 438, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that

comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID NO: 1 - 438 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO: 1 - 438 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences that flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding a nucleic acid disclosed herein (*e.g.*, SEQ ID NO: 1 - 438, antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of an mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of an mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of an mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil,

5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic

acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α -anomeric nucleic acid molecule. An α -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual α -units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave mRNA transcripts to thereby inhibit translation of an mRNA. A ribozyme having specificity for a nucleic acid of the invention can be designed based upon the nucleotide sequence of a DNA disclosed herein (*i.e.*, SEQ ID NO: 1 - 438). For example, a derivative of Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (*e.g.*, promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991) *Anticancer Drug Des.* 6: 569-84; Helene. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14: 807-15.

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see 5 Hyrup *et al.* (1996) *Bioorg Med Chem* 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. 10 The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996) above; Perry-O'Keefe *et al.* (1996) *PNAS* 93: 14670-675.

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific 15 modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization 20 (Hyrup *et al.* (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that 25 may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup 30 (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a

DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, *Proc. Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (*e.g.*, by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous

promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No.

5 WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., *ada*, *dhfr*, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection
10 methods results in co-amplification of the desired protein coding sequences in the cells.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or
15 electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the
20 present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or
25 other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor, New York (1989),
30 the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey
5 COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a
10 suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides
15 and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in
20 expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains,
25 *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order
30 to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As
5 described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory
10 elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties
15 of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a
20 simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more
25 selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the
30 negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result

in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO: 1-438 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NOs: 1 - 438 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NOs: 1 - 438 or (b) polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO: 1-438 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO: 1-438 or the corresponding full length or mature protein; and "substantial equivalents" thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO: 1-438.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the

protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., *Bio/Technology* 10, 773-778 (1992) and in R. S. McDowell, et al., *J. Amer. Chem. Soc.* 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, *e.g.*, pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (*e.g.*, an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein

activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, *e.g.*, Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag (1994); Sambrook, et al.,

in *Molecular Cloning: A Laboratory Manual*; Ausubel et al., *Current Protocols in Molecular Biology*. Polypeptide fragments that retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

5 The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for e.g., small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models
10 that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

 In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds
15 that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO: 1-438.

 The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the
20 protein.

 The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications
25 of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art
30 (see, e.g., U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that

are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, Calif., U.S.A. (the MaxBat™ kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (*i.e.*, from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl™ or Cibacrom blue 3GA Sepharose™; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as

those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, *e.g.*, silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, *e.g.*, targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, *e.g.*, antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, *etc.*, as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

30

4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobicity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus, or to the middle.

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.

5 In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprise one or more domains fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction
15 may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e.g.*, cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a
20 ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation,
25 restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to
30 complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for

example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the
5 fusion moiety is linked in-frame to the protein of the invention.

4.8 GENE THERAPY

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to
10 restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected *ex vivo*, *in situ*, or *in vivo* by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (e.g.,
15 liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present
20 invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes. Alternatively, it is contemplated that in other human
25 disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of
30 antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the

present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered *in vivo* to express the polynucleotides of the invention, wherein such polynucleotides are in operative
5 association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be
10 modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650,
15 PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., *ada*, *dhfr*, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired
20 protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the
25 endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional
30 initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein

produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (*gpt*) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.9 TRANSGENIC ANIMALS

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the

regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No.

- 5 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT
10 Publication No. WO94/28122, incorporated herein by reference.

- Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or
15 even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

4.10 USES AND BIOLOGICAL ACTIVITY

- 20 The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or
25 vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including
30 recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and

truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins;

5 chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

10 The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

4.10.1 RESEARCH USES AND UTILITIES

The polynucleotides provided by the present invention can be used by the

15 research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when

20 labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and

25 making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand

30 interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify

polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

4.10.2 NUTRITIONAL USES

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

Assays for T-cell or thymocyte proliferation include without limitation those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., *J. Immunol.* 137:3494-3500, 1986; Bertagnolli et al., *J. Immunol.* 145:1706-1712, 1990; Bertagnolli et al., *Cellular Immunology* 133:327-341, 1991; Bertagnolli, et al., *I. Immunol.* 149:3778-3783, 1992; Bowman et al., *I. Immunol.* 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In *Current Protocols in Immunology*. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin- γ , Schreiber, R. D. In *Current Protocols in Immunology*. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: *Measurement of Human and Murine*

- Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.
- Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

- A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells *in vivo* or *ex vivo* is expected to maintain and expand cell populations in a totipotent or pluripotent state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-

pharmaceuticals and the development of bio-sensors. The ability to produce large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other
5 neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or
10 cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-
15 CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration
20 of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder layer for the stem
25 cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for
30 generation of undifferentiated totipotent/pluripotent stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell

lines can also serve as a source of undifferentiated totipotent/pluripotent mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

5 Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be
10 useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host
15 rejection of replacement tissues after grafting or implantation.

 Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the
20 use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., *Differentiation*, 48: 173-182, (1991); Klug et al., *J. Clin. Invest.*, 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering* eds. Lanza et al., Academic Press (1997)).
25 Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the
30 invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell sources (including hematopoietic stem cells and embryonic stem cells) and

cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support e.g. as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders.

Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without
5 limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those
10 described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol
15 pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I.
20 Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

4.10.6 TISSUE GROWTH ACTIVITY

25 A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the
30 healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the

invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

5 A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase
10 activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally
15 formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue.
20 De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth
25 of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering
30 agent as a carrier as is well known in the art.

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

5 Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

10 **4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY**

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune
15 deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious
20 diseases caused by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the
25 treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitus,
30 myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may

also to be useful in the treatment of allergic reactions and conditions (*e.g.*, anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastborn et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxicol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue

transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., *Science* 257:789-792 (1992) and Turka et al., *Proc. Natl. Acad. Sci USA*, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul et al., *Fundamental Immunology*, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal

models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythematosus in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β_2 microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a

T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

10 Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

20 Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

30 Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte

Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

4.10.8 ACTIVIN/INHIBIN ACTIVITY

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be

useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., *Endocrinology* 91:562-572, 1972; Ling et al., *Nature* 321:779-782, 1986; Vale et al., *Nature* 321:776-779, 1986; Mason et al., *Nature* 318:659-663, 1985; Forage et al., *Proc. Natl. Acad. Sci. USA* 83:3091-3095, 1986.

4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell

population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

5 Therapeutic compositions of the invention can be used in the following:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement
10 and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J.
15 Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostasis or
20 thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving
25 or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those
30 described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al.,

Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991);
Schaub, Prostaglandins 35:467-474, 1988.

4.10.11 CANCER DIAGNOSIS AND THERAPY

5 Polypeptides of the invention may be involved in cancer cell generation,
proliferation or metastasis. Detection of the presence or amount of polynucleotides or
polypeptides of the invention may be useful for the diagnosis and/or prognosis of one or
more types of cancer. For example, the presence or increased expression of a
polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a
10 precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or
absence of the polypeptide may be associated with a cancer condition. Identification of
single nucleotide polymorphisms associated with cancer or a predisposition to cancer
may also be useful for diagnosis or prognosis.

Cancer treatments promote tumor regression by inhibiting tumor cell
15 proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to
support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or
invasiveness. Therapeutic compositions of the invention may be effective in adult and
pediatric oncology including in solid phase tumors/malignancies, locally advanced
tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases,
20 blood cell malignancies including multiple myeloma, acute and chronic leukemias, and
lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid
cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast
cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers
including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps
25 associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers
including bladder cancer and prostate cancer, malignancies of the female genital tract
including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in
the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers
including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas,
30 metastatic tumor cell invasion in the central nervous system, bone cancers including
osteomas, skin cancers including malignant melanoma, tumor progression of human skin

keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Kaposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine. Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguanzone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these

individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These *in vitro* models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wiley-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

4.10.12 RECEPTOR/LIGAND ACTIVITY

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley- Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static
5 conditions 7.28.1- 7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor
10 for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The
15 polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of
radioisotopes include, but are not limited to, tritium and carbon-14 . Examples of
20 colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

4.10.13 DRUG SCREENING

25 This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably
30 transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays.

Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well
5 known in the art.

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides
10 or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria
15 and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science* 282:63-68 (1998).

20 Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic
25 collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, *Curr. Opin. Biotechnol.* 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., *Mol. Biotechnol.* 9(3):205-23 (1998); Hruby et al., *Curr Opin Chem Biol.* 1(1):114-19 (1997); Dorner et al., *Bioorg Med Chem.* 4(5):709-15 (1996) (alkylated dipeptides).

30 Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the

“hit” to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules, that modulate (*i.e.*, increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2)

natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications i.e. phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

4.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease,

inflammation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic myelogenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

5 **4.10.16 LEUKEMIAS**

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia,
10 myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, *Medicine*, 2d Ed., J.B. Lippincott Co., Philadelphia).

15 **4.10.17 NERVOUS SYSTEM DISORDERS**

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of therapeutic utility, include but are not limited to nervous system injuries, and diseases
20 or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

- 25 (i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;
- (ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or
30 spinal cord infarction or ischemia;

(iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;

5 (iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;

10 (v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;

15 (vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;

(vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and

20 (viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

25 Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

- 30 (i) increased survival time of neurons in culture;
- (ii) increased sprouting of neurons in culture or in vivo;

(iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or

(iv) decreased symptoms of neuron dysfunction *in vivo*.

Such effects may be measured by any method known in the art. In preferred,
5 non-limiting embodiments, increased survival of neurons may be measured by the
method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting
of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol.
70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of
neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody
10 binding, Northern blot assay, *etc.*, depending on the molecule to be measured; and motor
neuron dysfunction may be measured by assessing the physical manifestation of motor
neuron disorder, *e.g.*, weakness, motor neuron conduction velocity, or functional
disability.

In specific embodiments, motor neuron disorders that may be treated according to
15 the invention include but are not limited to disorders such as infarction, infection,
exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may
affect motor neurons as well as other components of the nervous system, as well as
disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and
including but not limited to progressive spinal muscular atrophy, progressive bulbar
20 palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive
bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio
syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

4.10.18 OTHER ACTIVITIES

25 A polypeptide of the invention may also exhibit one or more of the following
additional activities or effects: inhibiting the growth, infection or function of, or killing,
infectious agents, including, without limitation, bacteria, viruses, fungi and other
parasites; effecting (suppressing or enhancing) bodily characteristics, including, without
limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue
30 pigmentation, or organ or body part size or shape (such as, for example, breast
augmentation or diminution, change in bone form or shape); effecting biorhythms or

circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without
5 limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related
10 diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

15

4.10.19 IDENTIFICATION OF POLYMORPHISMS

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential
20 predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of
25 this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an
30 appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which

appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

4.10.20 ARTHRITIS AND INFLAMMATION

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et al., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of

the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

5 **4.11 THERAPEUTIC METHODS**

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified
10 herein.

4.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a
15 disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary
20 according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01 μ g/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1 μ g/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable
25 parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

30

4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth factor (PDGF), transforming growth factors (TGF- α and TGF- β), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2,

anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

5 As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). Techniques for formulation and administration of the compounds of
10 the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, *e.g.*, treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions.
15 When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

20 In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as
25 treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the
30 attending physician will decide on the appropriate sequence of administering protein or

other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

4.12.1 ROUTES OF ADMINISTRATION

5 Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical
10 composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

 Alternately, one may administer the compound in a local rather than systemic
15 manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome
20 coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

 The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of
25 skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

30

4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an

isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives
5 known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

10 For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be
15 obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose,
20 hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc,
25 polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules
30 made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture

with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, *e.g.*, gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active

ingredient may be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, *e.g.* polyvinyl pyrrolidone; and other sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent.

Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional
5 strategies for protein or other active ingredient stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the
10 invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate,
15 potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their
20 surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with
25 co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a
30 liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist

in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 μ g to about 100 mg (preferably about 0.1 μ g to about 10 mg, more preferably about 0.1 μ g to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering

the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

5 The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and
10 polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the
15 above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging
20 from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose,
25 ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful
30 herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer

matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with
5 other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF- α and TGF- β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary
10 applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, *e.g.*,
15 amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (*e.g.*, bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition
20 of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such
25 polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate
30 or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC_{50} as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD_{50} and ED_{50} . Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, *e.g.*, Fingl et

al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from *in vitro* data.

5 Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for
10 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01 µg/kg to 100 mg/kg of body weight daily,
15 with the preferred dose being about 0.1 µg/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the
20 manner of administration and the judgment of the prescribing physician.

4.12.4 PACKAGING

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The
25 pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

30

4.13 ANTIBODIES

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, i.e., molecules that contain an antigen-binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab} , F_{ab}' and $F_{(ab)2}$ fragments, and an F_{ab} expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG₁, IgG₂, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, such as an amino acid sequence shown in SEQ ID NO: 1-438, and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues, or at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of alpha-2-macroglobulin-like protein that is located on the surface of the protein, *e.g.*, a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting

antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 5 1981, *Proc. Nat. Acad. Sci. USA* 78: 3824-3828; Kyte and Doolittle 1982, *J. Mol. Biol.* 157: 105-142, each of which is incorporated herein by reference in its entirety.

Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog 10 thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

The term "specific for" indicates that the variable regions of the antibodies of the invention recognize and bind polypeptides of the invention exclusively (*i.e.*, able to distinguish the polypeptide of the invention from other similar polypeptides despite 15 sequence identity, homology, or similarity found in the family of polypeptides), but may also interact with other proteins (for example, *S. aureus* protein A or other antibodies in ELISA techniques) through interactions with sequences outside the variable region of the antibodies, and in particular, in the constant region of the molecule. Screening assays to determine binding specificity of an antibody of the invention are well known and 20 routinely practiced in the art. For a comprehensive discussion of such assays, see Harlow et al. (Eds), *Antibodies A Laboratory Manual*; Cold Spring Harbor Laboratory; Cold Spring Harbor, NY (1988), Chapter 6. Antibodies that recognize and bind fragments of the polypeptides of the invention are also contemplated, provided that the antibodies are first and foremost specific for, as defined above, full-length polypeptides of the 25 invention. As with antibodies that are specific for full length polypeptides of the invention, antibodies of the invention that recognize fragments are those which can distinguish polypeptides from the same family of polypeptides despite inherent sequence identity, homology, or similarity found in the family of proteins.

Antibodies of the invention are useful for, for example, therapeutic purposes (by 30 modulating activity of a polypeptide of the invention), diagnostic purposes to detect or quantitate a polypeptide of the invention, as well as purification of a polypeptide of the

invention. Kits comprising an antibody of the invention for any of the purposes described herein are also comprehended. In general, a kit of the invention also includes a control antigen for which the antibody is immunospecific. The invention further provides a hybridoma that produces an antibody according to the invention. Antibodies of the
5 invention are useful for detection and/or purification of the polypeptides of the invention.

Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where
10 abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

The labeled antibodies of the present invention can be used for *in vitro*, *in vivo*,
15 and *in situ* assays to identify cells or tissues in which a fragment of the polypeptide of interest is expressed. The antibodies may also be used directly in therapies or other diagnostics. The present invention further provides the above-described antibodies immobilized on a solid support. Examples of such solid supports include plastics such as polycarbonate, complex carbohydrates such as agarose and Sepharose®, acrylic resins
20 and such as polyacrylamide and latex beads. Techniques for coupling antibodies to such solid supports are well known in the art (Weir, D.M. et al., "Handbook of Experimental Immunology" 4th Ed., Blackwell Scientific Publications, Oxford, England, Chapter 10 (1986); Jacoby, W.D. et al., Meth. Enzym. 34 Academic Press, N.Y. (1974)). The immobilized antibodies of the present invention can be used for *in vitro*, *in vivo*, and *in*
25 *situ* assays as well as for immuno-affinity purification of the proteins of the present invention.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example,
30 Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor

Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

4.13.1 POLYCLONAL ANTIBODIES

5 For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the
10 immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various
15 adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface-active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and *Corynebacterium parvum*, or similar immunostimulatory agents. Additional examples of
20 adjuvants that can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide
25 primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (*The Scientist*, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No.
30 8 (April 17, 2000), pp. 25-28).

4.13.2 MONOCLONAL ANTIBODIES

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen-binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are

sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also
5 have been described for the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen.
10 Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and
15 Pollard, Anal. Biochem., 107:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640
20 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography,
25 gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically
30 to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can

be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

4.13.3 HUMANIZED ANTIBODIES

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeven et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539). In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues that are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are

those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

5

4.13.4 HUMAN ANTIBODIES

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein. Human monoclonal antibodies can be prepared by the trioma
10 technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 *Immunol Today* 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the
15 practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. *Proc Natl Acad Sci USA* 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques,
20 including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely
25 resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (Bio/Technology 10, 779-783 (1992)); Lonberg et al. (Nature 368 856-859 (1994)); Morrison (Nature 368, 812-13 (1994)); Fishwild et al, (Nature Biotechnology 14, 845-51
30 (1996)); Neuberger (Nature Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals that are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the XenomouseTM as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells that secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in

culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

5 In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

10 4.13.5 FAB FRAGMENTS AND SINGLE CHAIN ANTIBODIES

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab)2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab)2}$ fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

20 4.13.6 BISPECIFIC ANTIBODIES

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have

30

different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh *et al.*, Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers that are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')₂ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example,

bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate $F(ab')_2$ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from *E. coli* and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody $F(ab')_2$ molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H

domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., *J. Immunol.* 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example,
5 trispecific antibodies can be prepared. Tutt et al., *J. Immunol.* 147:60 (1991).

Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3,
10 CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as
15 EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

4.13.7 HETEROCONJUGATE ANTIBODIES

Heteroconjugate antibodies are also within the scope of the present invention.
20 Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving
25 crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

30 4.13.8 EFFECTOR FUNCTION ENGINEERING

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus
5 generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., *J. Exp Med.*, 176: 1191-1195 (1992) and Shopes, *J. Immunol.*, 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. *Cancer Research*, 53:
10 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., *Anti-Cancer Drug Design*, 3: 219-230 (1989).

4.13.9 IMMUNOCONJUGATES

15 The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have
20 been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcun, 25 crotin, *sapaonaria officinalis* inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi , ^{131}I , ^{131}In , ^{90}Y , and ^{186}Re .

Conjugates of the antibody and cytotoxic agent are made using a variety of
30 bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as

dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active
5 fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

10 In another embodiment, the antibody can be conjugated to a "receptor" (such as streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

15

4.14 COMPUTER READABLE SEQUENCES

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer
20 readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled artisan can readily
25 appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer
30 readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NOs: 1 - 438 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NOs: 1 - 438 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., *J. Mol. Biol.* 215:403-410 (1990)) and BLAZE (Brutlag et al., *Comp. Chem.* 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present

invention comprise a data storage means having stored therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means.

Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention.

Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDE). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic

acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

4.15 TRIPLE HELIX FORMATION

5 In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA. Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in
10 transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while
15 antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

20 4.16 DIAGNOSTIC ASSAYS AND KITS

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

25 In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization
30 conditions with nucleic acid primers that anneal to a polynucleotide of the invention

under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., *An Introduction to Radioimmunoassay and Related Techniques*, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., *Techniques in Immunocytochemistry*, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., *Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology*, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or

antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in
5 separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a
10 container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the
15 enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

20 4.17 MEDICAL IMAGING

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such
25 methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide *in vivo* at the target site.

4.18 SCREENING ASSAYS

30 Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a

polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NOs: 1 - 438, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

- (a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and
- (b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

- 5 For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular
- 10 protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and
- 15 Kaspiczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

- In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a
- 20 skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or
- 25 can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

- Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456
- 30 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of

Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences
5 of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to
10 generate a pharmaceutical composition.

4.19 USE OF NUCLEIC ACIDS AS PROBES

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring
15 nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NOs: 1 - 438. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from any of the nucleotide sequences SEQ ID NOs: 1 - 438 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

20 Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection
25 of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to
30 synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled

nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

5 Fluorescent *in situ* hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

10

15

4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

Oligonucleotides, i.e., small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

20

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata *et al.*, 1985; Dahlen *et al.*, 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller *et al.*, 1988; 1989); all references being specifically incorporated herein.

25

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8)

30

3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources,
5 such as, e.g., Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed CovaLink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent
10 coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen *et al.*, (1991) Anal. Biochem. 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end
15 has been described (Rasmussen *et al.*, (1991). In this technology, a phosphoramidate bond is employed (Chu *et al.*, (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer
20 arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M
25 1-methylimidazole, pH 7.0 (1-MeIm₇), is then added to a final concentration of 10 mM 1-MeIm₇. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm₇, is made fresh and 25 ul added per well. The strips are
30 incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing

solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated
5 herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not
10 cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface,
15 as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

20 To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6,
25 incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies.
30 A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example,
5 Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of
10 DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook *et al.* (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990)
15 Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

20 One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, *CviJI*, described by Fitzgerald *et al.* (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

25 The restriction endonuclease *CviJI* normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme (*CviJI***), yield a quasi-random distribution of DNA fragments from the small molecule pUC19 (2688 base pairs). Fitzgerald *et al.* (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a *CviJI***
30 digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76

clones showed that *CviJI*** restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead
5 of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed)

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C.
10 The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

4.22 PREPARATION OF DNA ARRAYS

Arrays may be prepared by spotting DNA samples on a support such as a nylon
15 membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm², depending on the type of label used. By avoiding spotting in some preselected number of
20 rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate
25 (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane. Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm² and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville,
30 Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell

plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

5.0 EXAMPLES

5.1 EXAMPLE 1

Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems (ABI) sequencer to obtain the novel nucleic acid sequences. In

some cases RACE (Random Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction.

5.2 EXAMPLE 2

Novel Nucleic Acids

5 The novel nucleic acids of the present invention of the invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. The nucleic acids were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional
10 sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST version 119, gb pri 119, and UniGene version 119) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than
15 300 and percent identity greater than 95%.

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genbank (i.e., dbEST version
20 120, gb pri 120, UniGene version 120, Genpept release 120). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and cg-zip-2 (Hyseq, Inc.). The full-length nucleotide and amino acid sequences, including splice variants resulting from these procedures are shown in the Sequence Listing as SEQ ID NOS: 1- 438.

25 Table 1 shows the various tissue sources of SEQ ID NO: 1-438.

The nearest neighbor results for polypeptides encoded by SEQ ID NO: 1-438 were obtained by a BLASTP (version 2.0a1 19MP-WashU) search against Genpept, Geneseq and SwissProt databases using BLAST algorithm. The nearest neighbor result showed the closest homologue with functional annotation for SEQ ID NO: 1-438. The
30 translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The homologues with identifiable functions for SEQ ID NO: 1-

438 are shown in Table 2 below. Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), all the sequences were examined to determine whether they had identifiable signature regions. Table 3 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the Pfam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) polypeptides encoded by SEQ ID NO: 1-438 (i.e. SEQ ID NO: 1-438) were examined for domains with homology to certain peptide domains. Table 4 shows the name of the domain found, the description, the product of all the e-value of similar domains found, the pFam score for the identified domain within the sequence, number of similar domains found, and the position of the domain in the SEQ ID NO: being interrogated.

The GeneAtlas™ software package (Molecular Simulations Inc. (MSI), San Diego, CA) was used to predict the three-dimensional structure models for the polypeptides encoded by SEQ ID NO: 1-438 (i.e. SEQ ID NO: 1-438). Models were generated by (1) PSI-BLAST which is a multiple alignment sequence profile-based searching developed by Altschul et al, (Nucl. Acids. Res. 25, 3389-3408 (1997)), (2) High Throughput Modeling (HTM) (Molecular Simulations Inc. (MSI) San Diego, CA,) which is an automated sequence and structure searching procedure (<http://www.msi.com/>), and (3) SeqFold™ which is a fold recognition method described by Fischer and Eisenberg (J. Mol. Biol. 209, 779-791 (1998)). This analysis was carried out, in part, by comparing the polypeptides of the invention with the known NMR (nuclear magnetic resonance) and x-ray crystal three-dimensional structures as templates. Table 5 shows, "PDB ID", the Protein DataBase (PDB) identifier given to template structure; "Chain ID", identifier of the subcomponent of the PDB template structure; "Compound Information", information of the PDB template structure and/or its subcomponents; "PDB Function Annotation" gives function of the PDB template as annotated by the PDB files (<http://www.rcsb.org/PDB/>); start and end amino acid position of the protein sequence aligned; PSI-BLAST score, the verify score, the SeqFold score, and the Potential(s) of Mean Force (PMF). The verify score is produced by GeneAtlas™

software (MSI), is based on Dr. Eisenberg's Profile-3D threading program developed in Dr. David Eisenberg's laboratory (US patent no. 5,436,850 and Luthy, Bowie, and Eisenberg, Nature, 356:83-85 (1992)) and a publication by R. Sanchez and A. Sali, Proc. Natl. Acad. Sci. USA, 95:13597-12502. The verify score produced by GeneAtlas
 5 normalizes the verify score for proteins with different lengths so that a unified cutoff can be used to select good models as follows:

$$\text{Verify score (normalized)} = (\text{raw score} - 1/2 \text{ high score}) / (1/2 \text{ high score})$$

10 The PFM score, produced by GeneAtlas™ software (MSI), is a composite scoring function that depends in part on the compactness of the model, sequence identity in the alignment used to build the model, pairwise and surface mean force potentials (MFP). As given in Table 5, a verify score between 0 to 1.0, with 1 being the best, represents a good model. Similarly, a PMF score between 0 to 1.0, with 1 being the best, represents a good
 15 model. A SeqFold™ score of more than 50 is considered significant. A good model may also be determined by one of skill in the art based all the information in Table 5 taken in totality.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determined from using Neural Network SignalP
 20 V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering,
 25 Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et al reference, was obtained for the polypeptide sequences. Table 6 shows the position of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide. Table 7 correlates each of SEQ ID NO: 1-438 to a specific chromosomal location.

30 Table 8 is a correlation table of the novel polynucleotide sequences SEQ ID NO: 1-438, novel polypeptide sequences SEQ ID NO: 1-438, and their corresponding priority

nucleotide sequences in the priority application USSN 09/774,528, herein incorporated by reference in its entirety.

Table 1

Tissue Origin	RNA/Tissue Source	Library Name	SEQ ID NO:
adult brain	GIBCO	AB3001	76-77 91 106-107 115 134 163-164 178 203 232 255 265 276 279 322-323
adult brain	GIBCO	ABD003	16 19 24 77 80-81 85 89-90 92 96 98 105 110 116 121-123 125 130-132 134-136 138 142-143 151 153 158-159 163-164 184 191 193 196 198 200 208-209 213-214 216 219- 220 223 229 232-234 236 239 241 243 257- 259 262 265 267 274-276 278 284 292 302 317 321 324-325 327 337-338 340 348 359 371 391-392 400
adult brain	Clontech	ABR001	1 18-19 35 80 98 125 136 153 185 200 209 221 228-229 239 243 274-275 302 399-400
adult brain	Clontech	ABR0065	7-8 18 32 35 52 57 85 91 96 111 113 126 131 135 138-139 142 148 153-154 181 188 192 199 209-211 217 221 224 226 229 233 235 238 243 248 273 283-284 286 292 316 322 348 357 361 367 376 378 399 407 409 417 428
adult brain	Clontech	ABR008	2 4 6-11 19-21 23-25 31 35-37 39-41 45-46 72-73 76 80-81 85 88-90 94-95 97 102-105 109 111-112 114-119 121-122 126-131 134- 135 138-139 144 146-150 152-153 156-157 159 168-172 174-175 178 180 182 185-186 189-190 194 196 198-201 203 205-210 217 219 221-222 224 229-230 232-233 236-239 243-244 248 253-256 260-261 263-265 273 276 281-282 286-289 291-292 299-300 302 304 315-317 319 321-322 324 326 329 331- 332 341 352-357 360 362 365 367-368 370 376-377 379-380 383-384 387-389 391-392 394 396-402 407-410 412-413 419 425-426 433
adult brain	Clontech	ABR011	85 90
adult brain	BioChain	ABR012	148 213
adult brain	BioChain	ABR013	85 322
adult brain	Invitrogen	ABR014	9 23 85 146 200 233 282 321 330
adult brain	Invitrogen	ABR015	14 31 69 121 124 163 209 216 224 291 377
adult brain	Invitrogen	ABR016	92 136 219 279
adult brain	Invitrogen	ABT004	2 7-8 20-21 33 85 90-91 95 97 102-103 108 121 123 129-131 138-139 143 146 151 153 157-158 172 178 180 209-210 213 219 229- 230 232 234 239 308 321 330 360 365 370- 373 375 401 412
adipocytes	Stratagene	ADP001	3-4 23 36 79 81 106-107 116 129 133-134 147 151 154 158 179 181 192 196 222 230 256-257 287 292 297 313 329 359
adrenal gland	Clontech	ADR002	2 25 27 33 57 76 85-86 88 96 98 105-108 114 121-122 125 129-130 134 147 164 178 180 182 198-199 201 205 207-208 240-241 244 246 253-254 257 261 276 280 292 320 329 336 352 403
adult heart	GIBCO	AHR001	3 17-21 27 32 74 76 85 89-91 95-96 102-103 105-110 117 121 124-125 128 131 134-136 139 141 148 151-153 155-156 161 163 181- 182 186 190 193 198 200-201 205 207 211- 213 215 222

Table 1

Tissue Origin	RNA/Tissue Source	Library Name	SEQ ID NO:
			225 229-230 234 251-254 257-259 263 274-277 280 292-297 301 303-304 315-316 319 329-331 345 359 384 417 423-424
adult kidney	Invitrogen	AKT002	3 6 14 20-21 25-26 76 79 85 89 94 101 111 114 118 121 124 126 130-131 138 146 163 170 177-178 189 196 198 201 204 213 231 253-254 256-259 271 273-275 277 298 315 320 329 342
adult lung	GIBCO	ALG001	4 29 74 79 85 90 96 105 111 119 132 134 136 142 144 149 159 181 189 198 200 205-207 226 255 257 263 283 294 300 302-303 328 358-359 365 426
lymph node	Clontech	ALN001	6 16 31 105 120 215 257 295 306 309 359
young liver	GIBCO	ALV001	10-11 25-26 29 31 33 76 85 95 115 121-122 124 126 130 143 146 156 158 164 178 182 187 189 229 248 253-254 261 278 283 304 342 375
adult liver	Invitrogen	ALV002	10-12 23 26 31 33-34 38 53 56 90-92 94-95 118 121 124 128-129 138 141 146 148 153 156 161 171 178 198 216 232 248 253-254 256-257 264 302 306 365 375 383 396
adult liver	Clontech	ALV003	10-11 156 171 188
Ovary	Invitrogen	AOV001	3-8 10-11 14 16 19-22 24 27-31 34 36 57 73 75-76 81-82 85 89-91 94-98 104-109 111 115-116 121-128 130-131 134 136 138-139 141 143-144 146 149-150 152 155 157-160 163-166 170-173 175 177-178 180 182 184-187 189-190 193-194 196-197 200-201 212-213 215 217 222 225-226 228 230-233 235 241-243 245 248 253-259 261 266-267 270 272-273 276-278 283-285 287 289 292 297-299 305-306 315-317 319 323-325 329-331 341 343-344 352 358-359 363-366 382-383 386 389-390 412
Placenta	Invitrogen	APL001	73 92 117 135 182 194 232 246 261 272 282 359
placenta	Invitrogen	APL002	16 28 92 121 135 144 157 178 210 394
adult spleen	GIBCO	ASP001	3-4 16 32-33 35 90 96 99-100 123-125 128 131 134 136 139 151 178 181 189 194 200 210 218 229 251 253-255 257 276 283 307-309 315 329 354-355 357 392 400
testis	GIBCO	ATS001	22 73 82 91 96-97 104-105 117 124 130 134 164 173 200 209 222 233 241 253-254 257 285 287-288 305 325 329 351-353 359
bladder	Invitrogen	BLD001	4 108 130 150 212 226 236 240 242 257 276 287 305 395-396 415
bone marrow	Clontech	BMD001	1 4-5 22 29-30 34 72 85 88 90 92 94 98 104-107 109 111 113 117 120 123-125 128-129 132 135 140 142 144 146 152 163 165-166 170-173 177 180 182 186 189-190 198-209 215 222 225 232 240-246 251-252 260-261 273-275 277-280 283-285 300 316 318 346-347 359
bone marrow	GF	BMD002	1 4 7-8 10-11 16 19 25 31 49 61-62 72 74 76 80 85 88 90 93-95 97-101 109-110 112 114 116-117 121 126 129 132 135 141 144

Table 1

Tissue Origin	RNA/Tissue Source	Library Name	SEQ ID NO:
			146 149-150 154 157 160 162-163 165-166 170-172 175 178-180 182-183 186-190 192- 194 198-200 203 208 210-213 215 223 225 234 242 245 247 251-254 256-257 265 270 273 276-278 280 285 287 289 291 293-294 299 302 307 309 315 322 324 337-338 353 356-357 359 367 369 388 407 414 419 426 434
bone marrow	Clonetech	BMD007	144
*Mixture of 16 tissues - mRNA	VARIOUS VENDORS	CGd010	1 34-35 95 152 161 171 182 206 219 242 260 267 276 280 288 297 300 315-316 412
*Mixture of 16 tissues - mRNA	Various Vendors	CGd011	45 51 167 188 216 251-252
*Mixture of 16 tissues - mRNA	Various Vendors	CGd012	2 10-11 18-21 29 31 34-35 40 42-43 45 48 50-52 69-71 87-89 94-95 98-105 109 111-113 117 120 123 125 127 131 135-136 138 146 158 163 165-169 175 180 187-188 191 198 201 208 216 219-221 224 226 234 236 238- 239 241-246 251-252 260 264 270 276-277 279 281 283-284 287 295-296 314 319 321 327-328 331 333-334 337-341 343 351-352 361 365 369 379-380 387 389 395 397-399 402 406 410-412 417 419 424 426 431-433
*Mixture of 16 tissues - mRNA	Various Vendors	CGd013	29 48 101 146 167-169 187 219 234 327 333 339 341 365 412 433
*Mixture of 16 tissues - mRNA	Various Vendors	CGd015	29 86 90 95 98 110 113 118 132 158 171 184 193 218-220 243 284 310 385 410 419
*Mixture of 16 tissues - mRNA	Various Vendors	CGd016	3-4 20-21 29 38 85 88-89 95 105 119 122 131-133 140 185 211-212 225 256-257 273 276 302 318 379-380 390 400 419
colon	Invitrogen	CLN001	4 25 33 85 138 146 148 158-159 198 210 229 301 360 384 397
cervix	BioChain	CVX001	3 5 10-11 18 20-21 24-25 29 36 41 47 57 63 72 74 76 86 90 94 104 108-109 111 125 127 130 134 138 144 147 162 174 178-179 182 186 189 193 197 211 222 225-226 228 232 241 243 257 261 267 270 273-275 278-281 288-289 298 301-302 305 315 319 324-325 329 331 337-338 359 391-392 395 420
endothelial cells	Stratogene	EDT001	3-6 18-19 24 27-29 35 72 76 79-80 85 89 96 98 104-107 111 117 119-121 124-131 134 136 138-139 141 144 146-147 149 152 158-159 166-167 170-173 178-179 182-183 186-187 191 193-194 196-197 200 210-211 222-224 226 231-232 236 241 243 246 248 253-256 258-259 276 279 282 287 292 300 302-303 315 329 337-338 358-362 382-383 385-388
esophagus	BioChain	ESO002	257
fetal brain	Clontech	FBR001	34
fetal brain	Clontech	FBR004	3 139 144 271 284 337-338
fetal brain	Clontech	FBR006	4 6-11 14 18-21 24 28 31 37-38 40 63 76 85 87 89-90 94-95 97 105 108-109 112-113 115

Table 1

Tissue Origin	RNA/Tissue Source	Library Name	SEQ ID NO:
			117-120 127-130 133 138 140 144-146 148 170 172 175 180 182 186-188 190 192 194 199 201 203 209-210 215 219 222 229-230 232-233 240 243 245 253-255 270 273 276 281 288-289 292 295 304 315 317 319 324 330-331 356-357 359-360 364 367-368 379- 380 383 389 397 399-401 408-409 411 413 419 421 423
fetal brain	Invitrogen	FBT002	2 14 19 23 28 31 90 94 105 121 124 126 131 135 139 142 149 158 186 193 198 210 214- 215 232 239 242 248 255 267 326 332 365 369 371 376-383 394 399
fetal heart	Invitrogen	FHR001	4 7-8 10-11 14 17-21 28-29 31-32 60 64-65 73 85 87 92 95 102-103 105 108 111 113 117 119 121 125 128-129 134-135 141 152 154 156-157 160-161 172 176 178 194 196 198- 200 203 208 212 215 218 222 226 229 233- 234 253-257 261 265 272 276 281 292-293 295 303 305 319 325 327 337-338 341 345 349 354-355 367-368 389 395-396 398 412 417 436
fetal kidney	Clontech	FKD001	1 14 22 94 110 115 132 134-135 146 178 189 199 235-236 242 247 257 267 292 295 359
fetal kidney	Clontech	FKD002	22 31 38 40 46 94 122 127 131 156 160 194 198 229 253-254 270 292 303 319 354-355 389 396
fetal kidney	Invitrogen	FKD007	303
fetal lung	Clontech	FLG001	85 89 98-100 111 175 271 281 369
fetal lung	Invitrogen	FLG003	84 88 106-107 122 135 140 146 160 181 246 272 284 292 328 330 396 404 416 426
fetal liver-spleen	Soares	FLS001	1-3 6-12 14 19 23 28-31 33 57 59-60 72-76 78 80 83 85-138 140-141 143-144 146-155 157-161 163-197 200 204 208 210-211 223 225 230 232-233 235 241-243 245-266 268- 273 277 281 285-287 292 297 303 314 329 343 346-347 357-359 369 397 399 407 415
fetal liver-spleen	Soares	FLS002	1 3-4 6 10-12 23-24 29 31-33 35-37 53-54 74-76 79 81-82 86-89 91 94-95 99-104 106- 109 111-112 115 117-120 122 125-126 128- 129 132 134 136-138 141 146 149 153 157- 159 162-166 170 172 175 178-180 183 185- 191 194 196-197 205 207-212 222-225 228 232-233 239-241 248 251-252 255-256 258- 259 261-262 264 266-267 270-271 273-275 277-278 283 285 287 298 305 315 317-318 322 330-332 337-338 341 343 349 357-360 365 388 390-391 399 402 418 424
fetal liver-spleen	Soares	FLS003	12 29 91 98 111 119 156 163 165 178 186 193 210-211 276 286 315 322 346-347 357 365 424
fetal liver	Invitrogen	FLV001	7-8 14 35 118 122-123 129 146 182 211 230 232 248 251-252 264 287 304 337-338 344 346-347 352 365 367-369
fetal liver	Clontech	FLV002	102-103 147 149 300
fetal liver	Clontech	FLV004	73 85 105 108 118 122 126 141 156-157 161 165 170 178 180 182 194 215 218 225 240

Table 1

Tissue Origin	RNA/Tissue Source	Library Name	SEQ ID NO:
			242 247 251-252 292 330 337-338 369 407 411 440
fetal muscle	Invitrogen	FMS002	5 9 17-18 20-21 29 38 85 88 97 106-107 129 131 136 150-152 155 165 170 179 182 192- 193 212-213 229 234 242 258-259 270 282 286 289 300 316 319 345 351 354-355 360 389 396 408 410 437 439
fetal skin	Invitrogen	FSK001	2 4 7-8 29 33 42-43 49 51-52 58 74 82 85 90 94 110-111 116 118 121 133 136 138-139 145 151 154 156-157 161-162 172 181 184 186 193 198 200 205 207 209-211 222 227- 230 232 235 240 246 253-257 266 270 276 292 295 299 316 318 323 330 332 337-340 343 357 369 389 394-395 412 422 427
fetal skin	Invitrogen	FSK002	4 9 42 44 51 66 72 81 85 89-90 95 98 105 112-114 119 121 129 133 135 162 172 179- 182 197 200 208 210 231 243-244 272 304 316 330 339 354-355 357 360 389 395 410 417 437
fetal spleen	BioChain	FSP001	157 223
umbilical cord	BioChain	FUC001	4-6 20-21 25 29 73-74 83 87 89-91 94 101 109 120 123 125 128 130-131 133 141 143- 144 147 149 154 161 165 173 175 179 184 188 210-212 217 226 235 240 248 251-252 257 262 267 270 277 293 305 307 316 319 323 327 331 341 356 359 389 392 407 416
fetal brain	GIBCO	HFB001	2-4 16 20-21 74 77 85 89-91 96-98 104-105 111 114 118 121-122 124-125 127-128 131 134 137-140 142 144 146-148 151 153 158- 159 163-164 166 173 178 180 182 191 194 196 200 203 209-214 216-232 234-236 238- 239 243 253-255 263 270 272-273 276 281 292 310 316 319-321 332 348 357 359 365 399
macrophage	Invitrogen	HMP001	2 247
infant brain	Soares	IB2002	2-4 7-8 19-22 26-27 31-32 35 73-74 80 85 89 91 96-98 106-107 110 112 118-119 121- 122 125 128-131 134-144 148 153 164 166 172-173 177 180 186-187 191-194 196 202- 203 208-210 217 219 223-224 227 229 232- 234 236-237 239 241-243 245 248 253-259 273-275 278-279 282 287 294 298 309 314 317 322 327 330 333-334 341 348-350 360 368 376 379-380 382 396 406 424
infant brain	Soares	IB2003	3-4 20-21 26 28 31 35 73 85 95-96 110 113 119 122-123 130-131 135 138 140 142-143 146 153 155 170 172-173 186 191-193 196 209 219 223 226 229 233-234 236 239 245 248 253-254 256-257 273 279 291-292 304 314 337-338 343 359 367 371 376 397 413
lung, fibroblast	Strategene	LFB001	3 6 31 72-73 90 92 105-107 124 126-127 133 136 139 144 146 172 189 198 204 233 235 246 258-259 268 272 276 282 310 335 359 434
adult lung	Invitrogen	LGT002	4 19-21 28 33 35-36 49 72 79 81 85 88 90- 91 94-95 101 106-107 109 118 120-125 127

Table 1

Tissue Origin	RNA/Tissue Source	Library Name	SEQ ID NO:
			130-131 133 135-138 141-142 144 147 149 157 159-161 163 166 170-173 193-194 196- 197 212 216 218 221 223 226 228-229 231 233 241 247-248 253-255 257 261 266-267 270-275 277-278 282-283 292 298 301 303 315 318 324 331 335 354-355 359 367 369 381 392-393 398
leukocytes	GIBCO	LUC001	1-5 15 19-21 28 30-33 37 72 74 91 94-95 97-100 108-109 113 115 117 119-122 124-125 127-128 134-138 141 144 146-148 150-151 157-158 160 162-167 170-173 175-178 180-181 187 189 192 194 197 200 212-213 215-216 218-219 223 225 228-232 241-242 245-246 251-254 261 272-276 278-282 284 287-290 297-298 305 307 310-314 325 331 336 340 358-359 372 399 414
leukocytes	Clontech	LUC003	1 5 124 171 176 204 225 248 253-254 283 285 307 315
melanoma	Clontech	MEL004	4-5 24 37 72-74 81 85 106-107 113 136 177 203 205-207 209 231 243 284-285 315-316 320 326 359 374 428
mammary gland	Invitrogen	MMG001	2 4-5 7-8 10-12 29 31 34-35 38 50 80-81 85 89-90 92 94-97 105 108-109 119-124 126 128-130 135 138-139 141-142 144 146-147 153 155 157-159 163 178-179 181-182 198 200 209-210 219 223 228 230 232-233 235- 236 239 242 248 253-255 257 260-261 265- 267 270 272 281 287 292 294 315-316 318 324 327 330 337-340 354-355 357 369 372 383 392-395 401 404
neuron	Stratagene	NTD001	35 47 89-90 111 118 164 232 253-254 276 324 331 382
neuron	Stratagene	NTR001	20-21 37 122 147-149 170 179 181 186 212 226 258-259 265 276 369 436 438
neuronal cells	Stratagene	NTU001	7-8 37 55 80 85 112 118 126-127 133 138 140-141 151 170 181 210 214 225-226 236 243 287 328 330-331 357 383 400 436
pituitary gland	Clontech	PIT004	92 124 159 231
placenta	Clontech	PLA003	34 46 88 126 128 159 182 186 197 201 267 278 281-282 305 330 356 361 365 418
prostate	Clontech	PRT001	18 36 72 74 86 95 106-107 111 118 122 144 161 179 211 218 233 286 297
rectum	Invitrogen	REC001	9 31 85 121 128 147 171 200 219 257 292 340 394 398 407 412
salivary gland	Clontech	SAL001	3 24 38 80 122 136 147 189 241 282 296 310 351 392 395 415
saliva gland	Clontech	SALS03	118
small intestine	Clontech	SIN001	12 16 25 82-83 89-90 93 95 98 105-109 111 122-123 125-128 133-134 137 139 142 161 167 171 184 197 201 204 212 218 236 242- 243 248-249 253-254 257 267 276 284-285 292 297 300 303 310 313 317-318 325 340 343 352 354-355 359 383 391 416
spinal cord	Clontech	SPC001	3 39 84 86 94 96 105 115 117 130-131 134 136 141 143 148 155 176 190-191 203 213

Table 1

Tissue Origin	RNA/Tissue Source	Library Name	SEQ ID NO:
			224 233-234 236 239 279 283 298 320-321 332 336-338 356 359 365 404-406
thalamus	Clontech	THA002	2 20-21 23 74 81 85 105-106 116 121 131 146 171 185 188 200 209 219 233 239 256 258-259 273 276 362 399
thymus	Clontech	THM001	16 29 33 57 80 82 85 90 93-94 106-107 120 126 128 134 141 161 176 194 223 228 235 253-254 261 274-275 278 285 298 319 332 336 343 353 359 425
thymus	Clontech	THMc02	1-2 7-9 14 26 34 44 73 75 82 85 87 94 98 106-107 109-111 117 119-120 125-126 128- 129 139 141 144 147-148 151 154-155 162 165 170-172 175-176 179 182 186 193-194 199-200 208-209 213 218 233 235 240 242 247 253-254 257 265 276 281 287 290 305 307 312 319 336 342 354-356 359 364 367 399 408 412-413 415 419 421 426 429-433
thyroid gland	Clontech	THR001	3 5 7-8 28 30-31 33 73-77 80 82 85 88 90- 92 94 96-98 105-107 109 113 117 121-122 124-125 127-128 130 134 136 141 143 146- 148 152 161-163 166 175 177-178 181 194 199 201 204 210 212 216 218 223-226 228 230-231 234 236 241 243 246 253-257 261 270 272-273 276-278 281-283 287 292 295 298 303-304 308 315 323 329 335 352 359 362 401 416-417
trachea	Clontech	TRC001	88 138 180 226 228 279 359 411 436
uterus	Clontech	UTR001	3 10-11 23 77 92 106-107 109 111 141 197- 198 218 241 257 270 274-275 302 315 329 396 400 413

*The 16 tissue/mRNAs and their vendor sources are as follows: 1) Normal adult brain mRNA (Invitrogen), 2) Normal adult kidney mRNA (Invitrogen), 3) Normal fetal brain mRNA (Invitrogen), 4) Normal adult liver mRNA (Invitrogen), 5) Normal fetal kidney mRNA (Invitrogen), 6) Normal fetal liver mRNA (Invitrogen), 7) normal fetal skin mRNA (Invitrogen), 8) human adrenal gland mRNA (Clontech), 9) Human bone marrow mRNA (Clontech), 10) Human leukemia lymphoblastic mRNA (Clontech), 11) Human thymus mRNA (Clontech), 12) human lymph node mRNA (Clontech), 13) human spinal cord mRNA (Clontech), 14) human thyroid mRNA (Clontech), 15) human esophagus mRNA (BioChain), 16) human conceptional umbilical cord mRNA (BioChain).

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
1	gi9837125	Homo sapiens	membrane-associated nucleic acid binding protein mRNA, partial cds.	2553	54
1	gi7020305	Homo sapiens	cDNA FLJ20301 fis, clone HEP06569.	1728	47
1	gi7294120	Drosophila melanogaster	CG16807 gene product	1535	53
2	AAY57911	Homo sapiens	Human transmembrane protein HTMPN-35.	1258	82
2	AAB88406	Homo sapiens	Human membrane or secretory protein clone PSEC0162.	265	39
2	gi14272664	Homo sapiens	unnamed protein product	265	39
3	gi12654575	Homo sapiens	Similar to gp25L2 protein, clone MGC:2142 IMAGE:2967520, mRNA, complete cds.	1116	100
3	gi12845568	Mus musculus	putative	1099	98
3	gi996057	Homo sapiens	H.sapiens mRNA for gp25L2 protein.	1096	98
4	gi9971050	Homo sapiens	Human DNA sequence from clone RP11-526K24 on chromosome 20. Contains a novel gene, the 5' end of a novel gene, two CpG islands, ESTs, GSSs and STSs, complete sequence.	4348	99
4	AAB95086	Homo sapiens	Human protein sequence SEQ ID NO:16999.	3034	99
4	gi10433753	Homo sapiens	cDNA FLJ12307 fis, clone MAMMA1001908.	3034	99
5	gi4689106	Homo sapiens	NADH-ubiquinone oxidoreductase B8 subunit	505	100
5	gi2909862	Homo sapiens	NADH-ubiquinone oxidoreductase subunit CI-B8 mRNA, complete cds.	505	100
5	gi12539408	Homo sapiens	NDUFA2 gene for NADH dehydrogenase (ubiquinone) 1 alpha subcomplex 2, complete cds.	505	100
6	AAG64416	Homo sapiens	Human nucleoprotein.	3765	100
6	gi10443046	Homo sapiens	Human DNA sequence from clone RP11-465L10 on chromosome 20. Contains 10 CpG islands, ESTs, STSs and GSSs. Contains the gene for a novel protein similar to Drosophila CG11399, the gene for a novel C2H2 type zinc finger protein similar to chicken FZF-1, a Ferritin light polypeptide (FTL) pseudogene, the MMP9 gene for matrix metalloproteinase 9 (gelatinase B, 92kD gelatinase, 92kD type IV collagenase) (CLG4B), a novel gene, the SLC12A5 gene for solute carrier family 12, (potassium-chloride transporter) member 5 (KIAA1176) and the 3' end of gene KIAA1637, complete sequence.	3765	100
6	gi15426514	Homo sapiens	clone MGC:16205 IMAGE:3640928, mRNA, complete cds.	3765	100
7	AAG64416	Homo sapiens	Human nucleoprotein.	3366	100

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
7	gi10443046	Homo sapiens	Human DNA sequence from clone RP11-465L10 on chromosome 20. Contains 10 CpG islands, ESTs, STSs and GSSs. Contains the gene for a novel protein similar to Drosophila CG11399, the gene for a novel C2H2 type zinc finger protein similar to chicken FZF-1, a Ferritin light polypeptide (FTL) pseudogene, the MMP9 gene for matrix metalloproteinase 9 (gelatinase B, 92kD gelatinase, 92kD type IV collagenase) (CLG4B), a novel gene, the SLC12A5 gene for solute carrier family 12, (potassium-chloride transporter) member 5 (KIAA1176) and the 3' end of gene KIAA1637, complete sequence.	3366	100
7	gi15426514	Homo sapiens	clone MGC:16205 IMAGE:3640928, mRNA, complete cds.	3366	100
8	gi14571904	Rattus norvegicus	lysosomal amino acid transporter 1	2145	85
8	AAE04910	Homo sapiens	Human transporter and ion channel-23 (TRICH-23) protein.	1239	56
8	gi7297404	Drosophila melanogaster	CG13384 gene product	837	43
9	AAB73686	Homo sapiens	Human oxidoreductase protein ORP-19.	1301	98
9	gi7291405	Drosophila melanogaster	T3dh gene product	808	59
9	gi5824752	Caenorhabditis elegans	predicted using Genefinder~contains similarity to Pfam domain: PF00465 (Iron-containing alcohol dehydrogenases), Score=177.7, E-value=1.9e-50, N=2~cDNA EST EMBL:Z14517 comes from this gene; cDNA EST yk18d4.3 comes from this gene~cDNA EST yk18d4.5 comes from this gene; cDNA EST yk116f5.5 comes from this gene~cDNA EST yk132h3.3 comes from this gene; cDNA EST yk73d10.3 comes from this gene~cDNA EST yk93e9.3 comes from this gene; cDNA EST yk132h3.5 comes from this gene~cDNA EST yk73d10.5 comes from this gene; cDNA EST yk93e9.5 comes from this gene~cDNA EST yk135b6.5 comes from this gene; cDNA EST yk135b6.3 comes from this gene~cDNA EST yk201e5.3 comes from this gene; cDNA EST yk268b1.3 comes from this gene~cDNA EST yk261d6.3 comes	685	52

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			from this gene; cDNA EST yk262h1.3 comes from this gene~cDNA EST yk292h1.3 comes from this gene; cDNA EST yk304d8.3 comes from this gene~cDNA EST yk344b7.3 comes from this gene; cDNA EST yk351a6.3 comes from this gene~cDNA EST yk366d9.3 comes from this gene; cDNA EST yk368e3.3 comes from this gene~cDNA EST yk372c1.3 comes from this gene; cDNA EST yk389g3.3 comes from this gene~cDNA EST yk422d2.3 comes from this gene; cDNA EST yk381d7.3 comes from this gene~cDNA EST yk201e5.5 comes from this gene; cDNA EST yk267f6.5 comes from this gene~cDNA EST yk268b1.5 comes from this gene; cDNA EST yk261d6.5 comes from this gene~cDNA EST yk262h1.5 comes from this gene; cDNA EST yk292h1.5 comes from this gene~cDNA EST yk304d8.5 comes from this gene; cDNA EST yk344b7.5 comes from this gene~cDNA EST yk368e3.5 comes from this gene; cDNA EST yk372c1.5 comes from this gene~cDNA EST yk351a6.5 comes from this gene; cDNA EST yk366d9.5 comes from this gene~cDNA EST yk389g3.5 comes from this gene; cDNA EST yk422d2.5 comes from this gene~cDNA EST yk560f4.3 comes from this gene; cDNA EST yk625h5.3 comes from this gene~cDNA EST yk381d7.5 comes from this gene; cDNA EST yk560f4.5 comes from this gene~cDNA EST yk625h5.5 comes from this gene		
10	AAB73686	Homo sapiens	Human oxidoreductase protein ORP-19.	1552	99
10	gi7291405	Drosophila melanogaster	T3dh gene product	891	56
10	gi5824752	Caenorhabditis elegans	predicted using Genefinder~contains similarity to Pfam domain: PF00465 (Iron-containing alcohol dehydrogenases), Score=177.7, E-value=1.9e-50, N=2~cDNA EST EMBL:Z14517 comes from this gene; cDNA EST yk18d4.3 comes from this gene~cDNA EST yk18d4.5 comes from this gene; cDNA EST yk116f5.5 comes from this gene~cDNA EST yk132h3.3 comes from this gene; cDNA EST yk73d10.3 comes from this	730	51

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			gene~cDNA EST yk93e9.3 comes from this gene; cDNA EST yk132h3.5 comes from this gene~cDNA EST yk73d10.5 comes from this gene; cDNA EST yk93e9.5 comes from this gene~cDNA EST yk135b6.5 comes from this gene; cDNA EST yk135b6.3 comes from this gene~cDNA EST yk201e5.3 comes from this gene; cDNA EST yk268b1.3 comes from this gene~cDNA EST yk261d6.3 comes from this gene; cDNA EST yk262h11.3 comes from this gene~cDNA EST yk292h11.3 comes from this gene; cDNA EST yk304d8.3 comes from this gene~cDNA EST yk344b7.3 comes from this gene; cDNA EST yk351a6.3 comes from this gene~cDNA EST yk366d9.3 comes from this gene; cDNA EST yk368e3.3 comes from this gene~cDNA EST yk372c11.3 comes from this gene; cDNA EST yk389g3.3 comes from this gene~cDNA EST yk422d2.3 comes from this gene; cDNA EST yk381d7.3 comes from this gene~cDNA EST yk201e5.5 comes from this gene; cDNA EST yk267f6.5 comes from this gene~cDNA EST yk268b1.5 comes from this gene; cDNA EST yk261d6.5 comes from this gene~cDNA EST yk262h11.5 comes from this gene; cDNA EST yk292h11.5 comes from this gene~cDNA EST yk304d8.5 comes from this gene; cDNA EST yk344b7.5 comes from this gene~cDNA EST yk368e3.5 comes from this gene; cDNA EST yk372c11.5 comes from this gene~cDNA EST yk351a6.5 comes from this gene; cDNA EST yk366d9.5 comes from this gene~cDNA EST yk389g3.5 comes from this gene; cDNA EST yk422d2.5 comes from this gene~cDNA EST yk560f4.3 comes from this gene; cDNA EST yk625h5.3 comes from this gene~cDNA EST yk381d7.5 comes from this gene; cDNA EST yk560f4.5 comes from this gene~cDNA EST yk625h5.5 comes from this gene		
11	AAB85166	Homo sapiens	Human Bcl-G1 polypeptide.	1598	87
11	gi14598300	Homo sapiens	unnamed protein product	1598	87
11	gi12584085	Homo sapiens	apoptosis regulator BCL-G long form (BCLG) mRNA, complete cds.	1598	87
12	gi15077865	Mus musculus	bullous pemphigoid antigen 1-b	1253	82

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
12	gi15077863	Mus musculus	bullous pemphigoid antigen 1-a	1253	82
12	gi6624582	Homo sapiens	Human DNA sequence from clone RP1-61B2 on chromosome 6p11.2-12.3 Contains isoforms 1 and 3 of BPAG1 (bullous pemphigoid antigen 1 (230/240kD), an exon of a gene similar to murine MACF cytoskeletal protein, STSs and GSSs, complete sequence.	733	99
13	gi3702270	Homo sapiens	chromosome 19, cosmid R31408, complete sequence.	887	93
13	gi401845	Homo sapiens	ribosomal protein L18a mRNA, complete cds.	887	93
13	gi13960144	Homo sapiens	ribosomal protein L18a, clone MGC:4476 IMAGE:2961519, mRNA, complete cds.	887	93
14	AAB59090	Homo sapiens	Breast and ovarian cancer associated antigen protein sequence SEQ ID 798.	496	80
14	AAB44129	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1574.	453	81
14	gi14198321	Mus musculus	ribosomal protein L31	453	81
15	gi5689465	Homo sapiens	mRNA for KIAA1064 protein, partial cds.	5643	100
15	gi4884368	Homo sapiens	mRNA; cDNA DKFZp586L1220 (from clone DKFZp586L1220); partial cds.	1628	100
15	gi13161145	Homo sapiens	zinc finger protein mRNA, complete cds.	369	36
16	gi5870832	Mus musculus	skm-BOP1	2494	94
16	gi5870834	Mus musculus	skm-BOP2	2397	91
16	gi1809322	Mus musculus	t-BOP	2285	93
17	gi13938126	Mus musculus	RIKEN cDNA 3732409C05 gene	2678	98
17	gi12852375	Mus musculus	putative	2678	98
17	gi7024433	Torpedo marmorata	male sterility protein 2-like protein	2307	80
18	AAB95482	Homo sapiens	Human protein sequence SEQ ID NO:18007.	1572	67
18	gi14042809	Homo sapiens	cDNA FLJ14932 fis, clone PLACE1009639.	1572	67
18	gi12053165	Homo sapiens	mRNA; cDNA DKFZp434K0427 (from clone DKFZp434K0427); complete cds.	1572	67
19	gi7243159	Homo sapiens	mRNA for KIAA1389 protein, partial cds.	7842	99
19	gi4151328	Homo sapiens	high-risk human papilloma viruses E6 oncoproteins targeted protein E6TP1 alpha mRNA, complete cds.	3777	53
19	gi4151330	Homo sapiens	high-risk human papilloma viruses E6 oncoproteins targeted protein E6TP1 beta mRNA, complete cds.	3768	53
20	gi7243159	Homo sapiens	mRNA for KIAA1389 protein, partial cds.	7714	98
20	gi4151328	Homo sapiens	high-risk human papilloma viruses E6 oncoproteins targeted protein E6TP1	3806	54

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			alpha mRNA, complete cds.		
20	gi4151330	Homo sapiens	high-risk human papilloma viruses E6 oncoproteins targeted protein E6TP1 beta mRNA, complete cds.	3797	53
21	AAB95328	Homo sapiens	Human protein sequence SEQ ID NO:17595.	753	61
21	AAB93757	Homo sapiens	Human protein sequence SEQ ID NO:13432.	753	61
21	AAB29657	Homo sapiens	Human membrane-associated protein HUMAP-14.	753	61
22	gi7673373	Homo sapiens	SCAN-related protein RAZ1 (RAZ1) mRNA, partial cds.	1104	100
22	AAG93274	Homo sapiens	Human protein HP10543.	900	100
22	AAB42846	Homo sapiens	Human ORFX ORF2610 polypeptide sequence SEQ ID NO:5220.	900	100
23	gi7242963	Homo sapiens	mRNA for KIAA1304 protein, partial cds.	5409	99
23	gi3413874	Homo sapiens	mRNA for KIAA0456 protein, partial cds.	3695	67
23	AAB30852	Homo sapiens	Amino acid sequence of human signal transduction protein SGT6-1.	3685	68
24	AAG64386	Homo sapiens	Human alcohol dehydrogenase 39.	1228	77
24	gi12861800	Mus musculus	putative	1083	66
24	gi3878713	Caenorhabditis elegans	weak similarity with quinone oxidoreductase, contains similarity to Pfam domain: PF00107 (Zinc-binding dehydrogenases), Score=-80.6, E-value=6.2e-06, N=1~cDNA EST yk164b4.5 comes from this gene~cDNA EST yk164b4.3 comes from this gene~cDNA EST yk264f3.5 comes from this gene	556	39
25	AAE02629	Homo sapiens	Human secreted protein Zalpa37.	2481	100
25	gi14536691	Homo sapiens	unnamed protein product	2481	100
25	AAAY99419	Homo sapiens	Human PRO1780 (UNQ842) amino acid sequence SEQ ID NO:282.	1960	77
26	gi6102869	Homo sapiens	mRNA; cDNA DKFZp434H1235 (from clone DKFZp434H1235); partial cds.	831	100
26	gi12853439	Mus musculus	putative	789	94
26	gi2198807	Gallus gallus	monocarboxylate transporter 3	505	29
27	gi7299069	Drosophila melanogaster	CG11755 gene product	205	34
27	gi3875367	Caenorhabditis elegans	contains 3 cysteine rich repeats	136	41
27	gi3249080	Arabidopsis thaliana	Contains similarity to MYB transcription factor isolog T01O24.1 gb 2288980 from A. thaliana BAC gb AC002335.	69	35
28	gi11041628	Homo sapiens	RPL6 gene for ribosomal protein L6, complete cds.	1207	98
28	gi433416	Homo sapiens	Human mRNA for DNA-binding protein, TAXREB107, complete cds.	1207	98

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
28	gi13278717	Homo sapiens	ribosomal protein L6, clone MGC:1635 IMAGE:2823733, mRNA, complete cds.	1207	98
29	AAG03810	Homo sapiens	Human secreted protein, SEQ ID NO: 7891.	845	100
29	gi186800	Homo sapiens	Human ribosomal protein L12 mRNA, complete cds.	845	100
29	gi14198333	Homo sapiens	ribosomal protein L12, clone MGC:9760 IMAGE:3855674, mRNA, complete cds.	845	100
30	AAB95051	Homo sapiens	Human protein sequence SEQ ID NO:16849.	2965	100
30	gi10433519	Homo sapiens	cDNA FLJ12118 fis, clone MAMMA1000085, weakly similar to PUTATIVE CYSTEINYL-TRNA SYNTHETASE C29E6.06C (EC 6.1.1.16).	2965	100
30	gi13938199	Homo sapiens	hypothetical protein FLJ12118, clone MGC:15044 IMAGE:2822557, mRNA, complete cds.	2959	99
31	gi12858123	Mus musculus	putative	2441	73
31	gi7959195	Homo sapiens	mRNA for KIAA1467 protein, partial cds.	2232	100
31	gi13278148	Mus musculus	Similar to RIKEN cDNA 8430419L09 gene	794	83
32	gi15530305	Homo sapiens	Similar to RIKEN cDNA 1700045I19 gene, clone MGC:2647 IMAGE:3509621, mRNA, complete cds.	1245	84
32	gi9858803	Mus musculus	Zfp228	512	47
32	AAG75629	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6393.	511	46
33	gi8101071	Homo sapiens	golgin-like protein (GLP) gene, complete cds.	312	46
33	gi8099669	Homo sapiens	golgin-like protein (GLP) mRNA, complete cds.	312	46
33	gi11037008	Human herpesvirus 8	latent nuclear antigen	245	40
34	gi437985	Canis familiaris	Rab12 protein	1071	99
34	gi206531	Rattus norvegicus	RAB12	995	96
34	gi12851149	Mus musculus	putative	819	96
35	gi13543689	Homo sapiens	Similar to RIKEN cDNA 4933405K01 gene, clone MGC:14799 IMAGE:4068454, mRNA, complete cds.	1077	96
35	gi12805373	Mus musculus	Unknown (protein for MGC:7298)	950	84
35	gi12855529	Mus musculus	putative	642	79
36	gi12697979	Homo sapiens	mRNA for KIAA1717 protein, partial cds.	1982	100
36	gi1651678	Synechocystis sp. PCC 6803	ORF_ID:slr1485~hypothetical protein	185	34

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
36	gi2739367	Arabidopsis thaliana	putative phosphatidylinositol-4-phosphate 5-kinase	153	28
37	gi3800892	Homo sapiens	neurexin III-alpha gene, partial cds.	1255	99
37	gi294602	Rattus norvegicus	neurexin III-alpha	1160	91
37	gi205716	Rattus norvegicus	neurexin II-alpha-a	561	50
38	gi10047315	Homo sapiens	mRNA for KIAA1619 protein, partial cds.	4447	99
38	gi8217424	Homo sapiens	Human DNA sequence from clone RP11-108L7 on chromosome 10. contains part of the gene for a novel Insulin-like growth factor binding type protein with Kazal-type serine protease inhibitor domain, the gene for a novel protein similar to rat tricarboxylate carrier, the gene for a novel PDZ (DHR, GLGF) domain protein, the gene for a novel protein similar to KIAA0552, KIAA0341 and Fugu hypothetical protein 2, the gene for a novel protein similar to Plasmodium POM1 and C. elegans F46G11.1, a putative novel gene, the SEMA4G gene for semaphorin 4G and a novel gene. Contains ESTs, STSs, GSSs and seven putative CpG islands, complete sequence.	4407	99
38	gi4836757	Mus musculus	semaphorin subclass 4 member G	4021	90
39	gi10438664	Homo sapiens	cDNA: FLJ22324 fis, clone HRC05551.	307	100
39	gi13559240	Homo sapiens	Human DNA sequence from clone RP5-842G6 on chromosome 20. Contains the 3' end of a novel gene, the 3' end of the gene for a novel protein similar to SEL1L (sel-1 (suppressor of lin-12, C.elegans)-like), ESTs, STSs and GSSs, complete sequence.	307	100
39	gi13543669	Homo sapiens	hypothetical protein FLJ22324, clone MGC:14701 IMAGE:4247211, mRNA, complete cds.	307	100
40	gi14595019	Homo sapiens	mRNA for keratin 6 irs (KRT6IRS gene).	2615	99
40	gi6092075	Mus musculus	type II cytokeratin	2414	91
40	gi15559584	Homo sapiens	Similar to keratin 6A, clone MGC:20671 IMAGE:3639270, mRNA, complete cds.	1468	57
41	gi12655452	Homo sapiens	mRNA for keratin associated protein 4.7 (KRTAP4.7 gene).	1157	86
41	gi12655464	Homo sapiens	partial mRNA for keratin associated protein 4.15 (KRTAP4.15 gene).	1090	88
41	gi12655462	Homo sapiens	mRNA for keratin associated protein 4.14 (KRTAP4.14 gene).	1063	84

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
42	gi553772	Homo sapiens	Human Tcr-C-delta gene, exons 1-4; Tcr-V-delta gene, exons 1-2; T-cell receptor alpha (Tcr-alpha) gene, J1-J61 segments; and Tcr-C-alpha gene, exons 1-4.	110	100
42	gi4379087	Homo sapiens	mRNA for TCR alpha variable region, patient AF31.	73	46
42	AAW40057	Homo sapiens	Cellular transcriptional factor p300.	71	42
43	gi15866589	Capsella rubella	hypothetical protein	97	30
43	gi3879045	Caenorhabditis elegans	R102.6	96	34
43	AAV56133	Homo sapiens	Human N-methyl-D-aspartate receptor 2 subunit SEQ ID NO:54.	94	52
44	gi13569345	Homo sapiens	pregnancy-associated plasma preproprotein-A2 mRNA, complete cds.	9839	99
44	gi10639043	Homo sapiens	mRNA for pregnancy-associated plasma protein-E (PAPPE gene).	8966	99
44	gi1142970	Homo sapiens	Human pregnancy-associated plasma protein-A preproform (PAPPA) mRNA, complete cds.	3856	45
45	gi12851017	Mus musculus	putative	578	83
45	gi4490653	Schizosaccharomyces pombe	profilin.	186	35
45	gi440266	Acanthamoeba castellanii	profilin I	166	34
46	gi1617480	Comamonas testosteroni	unknown	712	82
46	gi3046394	Ralstonia eutropha	phbF	563	66
46	gi6683782	Burkholderia sp. DSMZ 9242	unknown	560	61
47	gi9229934	Mus musculus	midnolin	2103	78
47	AAB56832	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1410.	912	71
47	gi15929300	Homo sapiens	Similar to midnolin, clone IMAGE:3958934, mRNA, partial cds.	907	100
48	gi13377624	Homo sapiens	calicin mRNA, complete cds.	3089	99
48	gi854100	Homo sapiens	H.sapiens mRNA for calicin (partial).	3076	99
48	gi853784	Bos taurus	calicin	2896	91
49	AAB68411	Homo sapiens	Amino acid sequence of a human NOV2 polypeptide.	2131	100
49	AAV99407	Homo sapiens	Human PRO1337 (UNQ692) amino acid sequence SEQ ID NO:236.	2101	99
49	AAB68414	Homo sapiens	Amino acid sequence of NOV2 polypeptide clone TA-cgAL132708 A.	2014	99
50	gi12082748	Mus musculus	T-box transcription factor TBX18	2972	93
50	gi5102617	Homo sapiens	Human DNA sequence from clone 33L1 on chromosome 6q14.1-15. Contains the gene for novel T-box (Brachyury) family protein. Contains	2634	100

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			ESTs, STSs, GSSs and two putative CpG islands, complete sequence.		
50	gi12849661	Mus musculus	putative	2223	96
51	gi12843048	Mus musculus	putative	339	72
51	gi6691626	Homo sapiens	RAGE mRNA for advanced glycation endproducts receptor, complete cds.	111	32
51	gi190846	Homo sapiens	Human receptor for advanced glycosylation end products (RAGE) mRNA, partial cds.	111	32
52	AAG71840	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1521.	1313	85
52	AAG71839	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1520.	1226	81
52	AAG71837	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1518.	1159	77
53	AAB94026	Homo sapiens	Human protein sequence SEQ ID NO:14163.	966	98
53	gi10433955	Homo sapiens	cDNA FLJ12457 fis, clone NT2RM1000666, weakly similar to DNA-BINDING PROTEIN A.	966	98
53	gi7295442	Drosophila melanogaster	CG17334 gene product	302	47
54	gi8980396	Homo sapiens	mRNA for T-cell antigen receptor-alpha, clone Pil-1a, partial.	566	97
54	gi2358063	Homo sapiens	T-cell receptor alpha delta locus from bases 752679 to 1000555 (section 4 of 5) of the Complete Nucleotide Sequence.	565	100
54	gi623149	Macaca mulatta	T-cell receptor alpha	512	85
55	gi2792496	Rattus norvegicus	tulip 2	2437	86
55	gi4884288	Homo sapiens	mRNA; cDNA DKFZp566D133 (from clone DKFZp566D133); partial cds.	1983	99
55	AAB41763	Homo sapiens	Human ORFX ORF1527 polypeptide sequence SEQ ID NO:3054.	1976	98
56	gi15524592	Homo sapiens	unnamed protein product	1033	52
56	gi537514	Homo sapiens	Human arylacetamide deacetylase mRNA, complete cds.	1033	52
56	AAB54079	Homo sapiens	Human pancreatic cancer antigen protein sequence SEQ ID NO:531.	1017	51
57	AAB33831	Homo sapiens	Human secreted protein BLAST search protein SEQ ID NO: 175.	149	35
57	gi1109682	Bos taurus	G-protein gamma-12 subunit	149	35
57	AAW09416	Homo sapiens	Human G protein gamma-7 subunit.	144	33
58	gi12082750	Mus musculus	T-box transcription factor TBX20	1469	93
58	gi9909810	Mus musculus	T-box transcription factor	1469	93
58	gi7229717	Danio rerio	H15-related T-box transcription factor hrT	1346	85
59	gi4185946	Human endogenous retrovirus K	gag protein	146	26
59	gi5802821	Homo sapiens	endogenous retrovirus HERV-K108,	146	26

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			complete sequence.		
59	gi5802814	Homo sapiens	endogenous retrovirus HERV-K103, complete sequence.	146	26
60	AAB94756	Homo sapiens	Human protein sequence SEQ ID NO:15815.	126	42
60	gi332612	Gibbon ape leukemia virus	pol polyprotein	113	50
60	gi3133302	Sus scrofa	pol protein	110	53
61	gi10121625	Gillichthys mirabilis	60S acidic ribosomal protein P1	127	81
61	AAB44012	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1457.	125	78
61	AAB43434	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:879.	125	78
62	AAB12585	Homo sapiens	Human T cell activating protein SEQ ID NO:4.	140	37
62	gi12805221	Mus musculus	lymphocyte antigen 6 complex	140	37
62	gi198924	Mus musculus	Ly-6A.2	140	37
63	gi6969165	Homo sapiens	Human DNA sequence from clone RP3-475N16 on chromosome 6p12.3-21.2. Contains the genes for CTG4A, pre-T cell receptor alpha, a novel protein similar to RPL7A (60S ribosomal protein L7A) and the 3' end of gene KIAA0240. Contains ESTs, STSs, GSSs and four putative CpG islands, complete sequence.	573	67
63	gi12841727	Mus musculus	putative	512	59
63	gi15293877	Ictalurus punctatus	ribosomal protein L7	314	38
64	gi181573	Homo sapiens	Human cytokeratin 8 (CK8) gene, complete cds.	1147	79
64	gi181400	Homo sapiens	Human cytokeratin 8 mRNA, complete cds.	1147	78
64	gi400416	Homo sapiens	H.sapiens KRT8 mRNA for keratin 8.	1147	79
65	gi13620887	Mus musculus	mitochondrial ribosomal protein S6	633	100
65	gi13620885	Homo sapiens	MRPS6 mRNA for mitochondrial ribosomal protein S6, partial cds.	565	85
65	gi14603226	Homo sapiens	clone MGC:19576 IMAGE:4304420, mRNA, complete cds.	565	85
66	gi13537119	Homo sapiens	mRNA for PAR-6 gamma, complete cds.	1956	100
66	gi8037909	Mus musculus	PAR6A	1490	76
66	gi9453884	Homo sapiens	mRNA for 16-5-5, partial cds.	1304	93
67	AAB95293	Homo sapiens	Human protein sequence SEQ ID NO:17517.	776	79
67	AAG81270	Homo sapiens	Human AFP protein sequence SEQ ID NO:58.	776	79
67	gi14035848	Homo sapiens	unnamed protein product	776	79
68	gi7020759	Homo sapiens	cDNA FLJ20565 fis, clone REC00542.	930	60
68	gi15216181	Homo sapiens	mRNA for putative 67-11-3 protein.	927	60
68	gi15930069	Homo sapiens	Similar to hypothetical protein FLJ20565, clone MGC:8850	917	60

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			IMAGE:3914396, mRNA, complete cds.		
69	gi3228237	Homo sapiens	UHS KerB gene.	810	72
69	gi200962	Mus musculus	serine 1 ultra high sulfur protein	755	69
69	gi32472	Homo sapiens	H.sapiens mRNA for high-sulphur keratin.	749	71
70	AAB92789	Homo sapiens	Human protein sequence SEQ ID NO:11284.	3518	100
70	gi7022420	Homo sapiens	cDNA FLJ10407 fis, clone NT2RM4000520.	3518	100
70	gi13111786	Homo sapiens	hypothetical protein FLJ10407, clone MGC:970 IMAGE:3509727, mRNA, complete cds.	3511	99
71	gi13325178	Homo sapiens	Similar to RIKEN cDNA 2210016F16 gene, clone MGC:10999 IMAGE:3638524, mRNA, complete cds.	856	100
71	gi7291278	Drosophila melanogaster	CG9752 gene product	744	43
71	gi2854153	Caenorhabditis elegans	Hypothetical protein C11D2.4	729	45
72	gi7020991	Homo sapiens	cDNA FLJ20718 fis, clone HEP17872.	3013	100
72	gi15680144	Homo sapiens	hypothetical protein FLJ20718, clone IMAGE:4577269, mRNA, partial cds.	2906	99
72	gi10801646	Macaca fascicularis	hypothetical protein	1097	99
73	AAG93290	Homo sapiens	Human protein HP10650.	1215	100
73	gi14587195	Homo sapiens	FAPP1-associated protein 1 (FASP1) mRNA, complete cds.	1215	100
73	gi8118225	Homo sapiens	chromosome 21 unknown mRNA.	1215	100
74	gi10436998	Homo sapiens	cDNA: FLJ21011 fis, clone CAE04289.	2522	100
74	gi15030282	Homo sapiens	clone MGC:16827 IMAGE:3855873, mRNA, complete cds.	2522	100
74	gi8570641	Homo sapiens	clone 133K02 unknown mRNA.	2514	99
75	gi6599255	Homo sapiens	mRNA; cDNA DKFZp434C0328 (from clone DKFZp434C0328).	1612	100
75	gi6330416	Homo sapiens	mRNA for KIAA1201 protein, partial cds.	554	38
75	AAB74726	Homo sapiens	Human membrane associated protein MEMAP-32.	496	35
76	gi7021059	Homo sapiens	cDNA FLJ20758 fis, clone HEP01508.	1450	100
76	AAW88552	Homo sapiens	Secreted protein encoded by gene 19 clone HSAVU34.	1429	100
76	gi15341707	Homo sapiens	clone MGC:19979 IMAGE:3939273, mRNA, complete cds.	1429	100
77	AAB95410	Homo sapiens	Human protein sequence SEQ ID NO:17796.	774	100
77	gi10435394	Homo sapiens	cDNA FLJ13391 fis, clone PLACE1001241.	774	100
77	gi10503974	Homo sapiens	clone SP24 unknown mRNA.	765	99
78	gi7020587	Homo sapiens	cDNA FLJ20467 fis, clone KAT06638.	737	100

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
78	AAB42883	Homo sapiens	Human ORFX ORF2647 polypeptide sequence SEQ ID NO:5294.	530	100
78	AAB56642	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1220.	530	100
79	AAW93948	Homo sapiens	Human regulatory molecule HRM-4 protein.	441	91
79	gi12852696	Mus musculus	putative	386	47
79	gi12751103	Homo sapiens	PNAS-129 mRNA, complete cds.	348	100
80	gi7243053	Homo sapiens	mRNA for KIAA1336 protein, partial cds.	3851	99
80	gi7292144	Drosophila melanogaster	CG2069 gene product	1634	44
80	gi1065457	Caenorhabditis elegans	C54G7.4 gene product	706	25
81	gi10439581	Homo sapiens	cDNA: FLJ23023 fis, clone LNG01678.	652	100
81	gi7021132	Homo sapiens	cDNA FLJ20813 fis, clone ADSE01247.	652	100
81	AAG74674	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:5438.	556	92
82	gi5262611	Homo sapiens	mRNA; cDNA DKFZp434I114 (from clone DKFZp434I114); complete cds.	838	100
82	gi11493368	Homo sapiens	Human DNA sequence from clone RP5-1009E24 on chromosome 20 Contains the SN gene encoding sialoadhesin, a novel gene similar to KIAA0417, the CENPB gene for centromere protein B, the CDC25B gene for Cell division cycle protein 25B, three novel genes, the 5' end of gene KIAA1271, nine CpG islands, ESTs, STSs and GSSs, complete sequence.	838	100
82	gi13543798	Mus musculus	RIKEN cDNA 4931426K16 gene	680	92
83	AAB57003	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1581.	1302	99
83	AAR60558	Homo sapiens	Human basigin I.	1302	99
83	gi3492872	Homo sapiens	chromosome 19, cosmid F18382 (LLNLF-140D2) and 3' overlapping restriction fragment, complete sequence.	1302	99
84	gi9187614	Homo sapiens	mRNA full length insert cDNA clone EUROIMAGE 1759349.	580	100
84	AAB01394	Homo sapiens	Neuron-associated protein.	70	39
84	AAB54358	Homo sapiens	Human pancreatic cancer antigen protein sequence SEQ ID NO:810.	70	39
85	gi15986445	Homo sapiens	p90 autoantigen mRNA, complete cds.	4513	99
85	gi7959315	Homo sapiens	mRNA for KIAA1524 protein, partial cds.	4357	99
85	AAB95207	Homo sapiens	Human protein sequence SEQ ID NO:17311.	2341	100
86	gi7959231	Homo sapiens	mRNA for KIAA1485 protein, partial cds.	5813	99

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
86	AAB40418	Homo sapiens	Human ORFX ORF182 polypeptide sequence SEQ ID NO:364.	708	99
86	gi5901529	Homo sapiens	C2H2 type Kruppel-like zinc finger protein splice variant b (ZNF236) mRNA, complete cds.	520	24
87	gi7243270	Homo sapiens	mRNA for KIAA1436 protein, partial cds.	4604	99
87	gi5051974	Mus musculus	F2 alpha prostoglandin regulatory protein	4195	89
87	gi1054884	Rattus norvegicus	prostaglandin F2a receptor regulatory protein precursor	4191	88
88	gi13241286	Mus musculus	GABA(A) receptor-associated protein-like 2	607	100
88	gi2104570	Rattus norvegicus	GEF-2	607	100
88	gi4433387	Bos taurus	general protein transport factor p16	607	100
89	gi15859535	Homo sapiens	unnamed protein product	5935	99
89	gi3043606	Homo sapiens	mRNA for KIAA0541 protein, partial cds.	5890	100
89	gi15624075	Homo sapiens	TGF-beta resistance-associated protein TRAG (TRAG) mRNA, partial cds.	5719	96
90	gi337370	Homo sapiens	Human rapamycin- and FK506-binding protein, complete cds.	740	100
90	gi13097252	Homo sapiens	Similar to FK506 binding protein 2 (13 kDa), clone MGC:5177 IMAGE:3445148, mRNA, complete cds.	740	100
90	AAQ31004_aa 1	Homo sapiens	hRFKBP cDNA.	735	99
91	gi12053147	Homo sapiens	mRNA; cDNA DKFZp434F1726 (from clone DKFZp434F1726).	1450	100
91	gi412195	Homo sapiens	unknown	265	98
91	AAR04931	Homo sapiens	Interferon-gamma receptor segment from clone 39 responsible for binding the target.	260	96
92	gi10437948	Homo sapiens	cDNA: FLJ21783 fis, clone HEP00284.	3276	100
92	AAB95352	Homo sapiens	Human protein sequence SEQ ID NO:17643.	1953	99
92	gi10435077	Homo sapiens	cDNA FLJ13171 fis, clone NT2RP3003819.	1953	99
93	gi12803319	Homo sapiens	clone MGC:3090 IMAGE:3347913, mRNA, complete cds.	4837	99
93	gi14044064	Homo sapiens	hypothetical protein DKFZp762M115, clone MGC:14418 IMAGE:4302613, mRNA, complete cds.	4831	99
93	gi10047337	Homo sapiens	mRNA for KIAA1630 protein, partial cds.	4671	100
94	AAB70535	Homo sapiens	Human PRO5 protein sequence SEQ ID NO:10.	2979	100
94	gi13185719	Homo sapiens	unnamed protein product	2979	100
94	AAB94106	Homo sapiens	Human protein sequence SEQ ID NO:14334.	2334	100
95	gi12837873	Mus musculus	putative	2370	75

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
95	gi13195574	Mus musculus	Praj1 isoform a	2339	75
95	AAB93847	Homo sapiens	Human protein sequence SEQ ID NO:13691.	1941	99
96	gi2224543	Homo sapiens	Human mRNA for KIAA0301 gene, partial cds.	10626	100
96	gi7529572	Homo sapiens	Human DNA sequence from clone RP1-12208 on chromosome 6q14.2-16.1. Contains the 3' part of a novel gene partially coded for by KIAA0301, a novel gene and the 3' part of the gene KIAA0957. Contains ESTs, STSs, GSSs and a putative CpG island, complete sequence.	10626	100
96	gi10727627	Drosophila melanogaster	CG13185 gene product	1452	34
97	AAB82318	Homo sapiens	Human immunoglobulin receptor IRTA5 protein.	2235	100
97	gi15528831	Homo sapiens	Fc receptor-like protein 1 (FCRH1) mRNA, complete cds.	2235	100
97	gi9930921	Homo sapiens	Human DNA sequence from clone RP11-367J7 on chromosome 1. Contains (part of) two or more genes for novel Immunoglobulin domains containing proteins, a SON DNA binding protein (SON) pseudogene, a voltage-dependent anion channel 1 (VDAC1) (plasmalemmal porin) pseudogene, ESTs, STSs and GSSs, complete sequence.	1533	100
98	AAB82318	Homo sapiens	Human immunoglobulin receptor IRTA5 protein.	2177	98
98	gi15528831	Homo sapiens	Fc receptor-like protein 1 (FCRH1) mRNA, complete cds.	2177	98
98	gi9930921	Homo sapiens	Human DNA sequence from clone RP11-367J7 on chromosome 1. Contains (part of) two or more genes for novel Immunoglobulin domains containing proteins, a SON DNA binding protein (SON) pseudogene, a voltage-dependent anion channel 1 (VDAC1) (plasmalemmal porin) pseudogene, ESTs, STSs and GSSs, complete sequence.	1533	100
99	gi10438861	Homo sapiens	cDNA: FLJ22461 fis, clone HRC10107.	4904	100
99	gi15079400	Homo sapiens	clone MGC:16796 IMAGE:3855477, mRNA, complete cds.	4899	99
99	AAU03497	Homo sapiens	Human sterol sensing domain protein.	4047	99
100	gi6524024	Mus musculus	mammalian inositol hexakisphosphate kinase 1	1031	50
100	gi10280996	Rattus norvegicus	inositol hexakisphosphate kinase	1027	49
100	gi6683115	Homo sapiens	mRNA for KIAA0263 protein, partial	1021	49

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			cds.		
101	gi6524024	Mus musculus	mammalian inositol hexakisphosphate kinase 1	1037	51
101	gi10280996	Rattus norvegicus	inositol hexakisphosphate kinase	1033	50
101	gi6683115	Homo sapiens	mRNA for KIAA0263 protein, partial cds.	1027	50
102	gi13623311	Homo sapiens	clone IMAGE:3948563, mRNA, partial cds.	1629	100
102	gi3135968	Homo sapiens	Human DNA sequence from clone XXbac-3418 on chromosome 6p21.3-22.1. Contains the 5' end of the ZNF184 gene for Kruppel-like zinc finger protein 184, a heterogeneous nuclear ribonucleoprotein A1 (HNRPA1) pseudogene, a CD83 antigen pseudogene, ESTs, STSs, GSSs and three CpG islands, complete sequence.	1627	47
102	gi1769491	Homo sapiens	Human kruppel-related zinc finger protein (ZNF184) mRNA, partial cds.	1625	47
103	gi16198398	Homo sapiens	clone MGC:27353 IMAGE:4671816, mRNA, complete cds.	2606	85
103	gi829151	Homo sapiens	H.sapiens ZNF37A mRNA for zinc finger protein.	1371	99
103	gi9801232	Homo sapiens	Human DNA sequence from clone RP11-508N22 on chromosome 10 Contains part of a novel gene (HSPC025), part of the ZNF37A (zinc finger protein 37a (KOX 21)) gene, part of a putative novel gene, ESTs, STSs, GSSs and a CpG Island, complete sequence.	1371	99
104	gi12053123	Homo sapiens	mRNA; cDNA DKFZp434K1421 (from clone DKFZp434K1421); complete cds.	2624	100
104	gi7292866	Drosophila melanogaster	CG15747 gene product	362	31
104	gi7549210	Babesia bigemina	200 kDa antigen p200	298	21
105	gi12053123	Homo sapiens	mRNA; cDNA DKFZp434K1421 (from clone DKFZp434K1421); complete cds.	2898	100
105	gi6841130	Homo sapiens	HSPC095 mRNA, partial cds.	419	100
105	gi7292866	Drosophila melanogaster	CG15747 gene product	364	30
106	gi10438207	Homo sapiens	cDNA: FLJ21977 fis, clone HEP05976.	1978	99
106	gi15012167	Homo sapiens	hypothetical protein FLJ21977, clone MGC:14918 IMAGE:3936410, mRNA, complete cds.	1974	99
106	AAB42499	Homo sapiens	Human ORFX ORF2263 polypeptide sequence SEQ ID NO:4526.	1392	100
107	gi1228035	Homo sapiens	Human mRNA for KIAA0191 gene,	8020	99

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			partial cds.		
107	gi12697967	Homo sapiens	mRNA for KIAA1711 protein, partial cds.	1593	58
107	AAB94636	Homo sapiens	Human protein sequence SEQ ID NO:15515.	1004	52
108	AAG81252	Homo sapiens	Human AFP protein sequence SEQ ID NO:22.	2146	99
108	gi14035812	Homo sapiens	unnamed protein product	2146	99
108	gi10440123	Homo sapiens	cDNA: FLJ23436 fis, clone HRC12692.	2054	100
109	gi200009	Mus musculus	myosin I	5386	96
109	gi1666471	Mus musculus	myosin I heavy chain	5360	94
109	gi56733	Rattus norvegicus	myosin I heavy chain	5268	91
110	gi12053045	Homo sapiens	mRNA; cDNA DKFZp434K1115 (from clone DKFZp434K1115); complete cds.	4840	100
110	AAB65631	Homo sapiens	Novel protein kinase, SEQ ID NO: 158.	4835	99
110	gi14133215	Homo sapiens	mRNA for KIAA0781 protein, partial cds.	4678	100
111	gi12642596	Homo sapiens	nuclear receptor co-repressor/HDAC3 complex subunit TBLR1 (TBLR1) mRNA, complete cds.	2725	100
111	AAB95225	Homo sapiens	Human protein sequence SEQ ID NO:17352.	2720	99
111	gi10434648	Homo sapiens	cDNA FLJ12894 fis, clone NT2RP2004170, moderately similar to Homo sapiens mRNA for transducin (beta) like 1 protein.	2720	99
112	gi2224557	Homo sapiens	Human mRNA for KIAA0308 gene, partial cds.	6666	99
112	AAAY23330	Homo sapiens	Human tumour suppressor (kismet) protein.	5759	98
112	gi7243213	Homo sapiens	mRNA for KIAA1416 protein, partial cds.	5264	59
113	gi12856019	Mus musculus	putative	1527	95
113	gi3947604	Caenorhabditis elegans	cDNA EST yk129f1.3 comes from this gene~cDNA EST yk129f1.5 comes from this gene~cDNA EST yk203e4.3 comes from this gene~cDNA EST yk191a9.3 comes from this gene~cDNA EST yk262c10.3 comes from this gene~cDNA EST yk278f9.3 comes from this gene~cDNA EST yk325c7.3 comes from this gene~cDNA EST yk337f1.3 comes from this gene~cDNA EST yk449a2.3 comes from this gene~cDNA EST yk203e4.5 comes from this gene~cDNA EST yk191a9.5 comes from this gene~cDNA EST yk278f9.5 comes from this gene~cDNA EST yk262c10.5 comes from this	787	41

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			gene~cDNA EST yk325c7.5 comes from this gene~cDNA EST yk337f1.5 comes from this gene~cDNA EST yk448g10.5 comes from this gene~cDNA EST yk449a2.5 comes from this gene~cDNA EST yk636e2.3 comes from this gene~cDNA EST yk636e2.5 comes from this gene~cDNA EST yk550e8.3 comes from this gene~cDNA EST yk557a9.3 comes from this gene~cDNA EST yk579c12.3 comes from this gene~cDNA EST yk614e7.3 comes from this gene~cDNA EST yk653f1.3 comes from this gene~cDNA EST yk672b2.3 comes from this gene~cDNA EST yk550e8.5 comes from this gene~cDNA EST yk556b1.5 comes from this gene~cDNA EST yk557a9.5 comes from this gene~cDNA EST yk579c12.5 comes from this gene~cDNA EST yk606c8.5 comes from this gene~cDNA EST yk614e7.5 comes from this gene		
113	gi3947603	Caenorhabditis elegans	cDNA EST yk167h7.3 comes from this gene~cDNA EST yk167h7.5 comes from this gene~cDNA EST yk289g5.3 comes from this gene~cDNA EST yk332h9.3 comes from this gene~cDNA EST yk289g5.5 comes from this gene~cDNA EST yk332h9.5 comes from this gene~cDNA EST yk391h4.5 comes from this gene~cDNA EST yk653f1.5 comes from this gene	787	41
114	gi9280136	Macaca fascicularis	unnamed protein product	3431	95
114	gi4262617	Caenorhabditis elegans	contains similarity to dual specificity phosphatase, catalytic domain (Pfam:PF00782, Score=16.8, E=7.4e-05, N=1)	470	35
114	gi5706724	Homo sapiens	Cdc14B3 phosphatase mRNA, complete cds.	166	30
115	AAB95254	Homo sapiens	Human protein sequence SEQ ID NO:17423.	3114	99
115	gi14042385	Homo sapiens	cDNA FLJ14693 fis, clone NT2RP2005360, weakly similar to Homo sapiens sentrin/SUMO-specific protease (SENPI) mRNA.	3114	99
115	gi10314023	Homo sapiens	sentrin-specific protease (SENPI2) mRNA, complete cds.	3107	99
116	gi4240227	Homo sapiens	mRNA for KIAA0869 protein, partial cds.	4417	98

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
116	gi13879506	Mus musculus	Unknown (protein for IMAGE:3963643)	4063	89
116	AAB93267	Homo sapiens	Human protein sequence SEQ ID NO:12300.	1895	97
117	gi13235092	Homo sapiens	mRNA for testis specific protein A14 (TSGA14 gene).	1957	100
117	gi10438839	Homo sapiens	cDNA: FLJ22445 fis, clone HRC09438.	1950	99
117	gi13235344	Mus musculus	testis specific protein a14	1704	87
118	gi7959279	Homo sapiens	mRNA for KIAA1509 protein, partial cds.	6769	99
118	AAB94101	Homo sapiens	Human protein sequence SEQ ID NO:14322.	1871	99
118	gi10434073	Homo sapiens	cDNA FLJ12531 fis, clone NT2RM4000199.	1871	99
119	AAM00936	Homo sapiens	Human bone marrow protein, SEQ ID NO: 412.	3350	100
119	AAB42828	Homo sapiens	Human ORFX ORF2592 polypeptide sequence SEQ ID NO:5184.	2064	100
119	gi9557949	Homo sapiens	mRNA for hypothetical protein (ORF1), clone Telethon(Italy_B41)_Strait02270_FL142.	1931	100
120	AAB11082	Homo sapiens	Human secreted protein ZALPHA13 protein.	2783	93
120	gi11230043	Homo sapiens	unnamed protein product	2783	93
120	AAB37988	Homo sapiens	Human secreted protein encoded by gene 5 clone HDPAS92.	2747	93
121	gi12852526	Mus musculus	putative	1689	80
121	AAB41765	Homo sapiens	Human ORFX ORF1529 polypeptide sequence SEQ ID NO:3058.	1576	100
121	gi4406663	Homo sapiens	clone 24945 mRNA sequence, partial cds.	1576	100
122	AAR22958	Homo sapiens	Human proteasome component HC5.	1010	85
122	gi220026	Homo sapiens	Human mRNA for proteasome subunit HC5.	1010	85
122	gi3790135	Homo sapiens	Human DNA sequence from clone RP1-191N21 on chromosome 6q27. Contains a 7 transmembrane receptor (rhodopsin family) (olfactory receptor like) pseudogene, the PDCD2 gene for programmed cell death 2 (RP8 homolog), the TBP gene for TATA box binding protein, the gene for proteasome subunit HC5, ESTs, STSs and GSSs, complete sequence.	1010	85
123	AAB21027	Homo sapiens	Human nucleic acid-binding protein, NuABP-31.	1456	100
123	AAB45146	Homo sapiens	Human secreted protein sequence encoded by gene 27 SEQ ID NO:87.	1456	100
123	gi4884258	Homo sapiens	mRNA; cDNA DKFZp564O092 (from clone DKFZp564O092); partial cds.	1430	100
124	gi13325436	Homo sapiens	Similar to RIKEN cDNA	1394	100

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			C330013D18 gene, clone MGC:11226 IMAGE:3937599, mRNA, complete cds.		
124	gi13559363	Homo sapiens	MRPL9 mRNA for mitochondrial ribosomal protein L9 (L9mt), complete cds.	1388	99
124	AAG93251	Homo sapiens	Human protein HP02612.	1153	86
125	AAB85507	Homo sapiens	Human protein kinase SGK164.	2949	100
125	gi13543922	Homo sapiens	Similar to RIKEN cDNA 5430416A05 gene, clone MGC:12903 IMAGE:3537086, mRNA, complete cds.	2913	100
125	gi12856491	Mus musculus	putative	2135	79
126	gi12653817	Homo sapiens	Similar to Male-specific RNA 84Dd, clone MGC:3092 IMAGE:3349383, mRNA, complete cds.	3399	100
126	AAB94115	Homo sapiens	Human protein sequence SEQ ID NO:14356.	3392	99
126	gi10434102	Homo sapiens	cDNA FLJ12549 fis, clone NT2RM4000689.	3392	99
127	gi7243187	Homo sapiens	mRNA for KIAA1403 protein, partial cds.	6448	98
127	gi12652971	Homo sapiens	clone MGC:858 IMAGE:3357380, mRNA, complete cds.	3992	100
127	AAB92872	Homo sapiens	Human protein sequence SEQ ID NO:11460.	3987	99
128	AAB94324	Homo sapiens	Human protein sequence SEQ ID NO:14807.	1779	99
128	gi10434528	Homo sapiens	cDNA FLJ12816 fis, clone NT2RP2002609, weakly similar to 2-HYDROXYMUCONIC SEMIALDEHYDE HYDROLASE (EC 3.1.1.-).	1779	99
128	AAB42143	Homo sapiens	Human ORFX ORF1907 polypeptide sequence SEQ ID NO:3814.	1521	100
129	gi6329945	Homo sapiens	mRNA for KIAA1140 protein, partial cds.	1857	52
129	gi12805043	Homo sapiens	clone IMAGE:3461487, mRNA, partial cds.	1279	54
129	gi7302173	Drosophila melanogaster	BcDNA:LD21719 gene product	1261	35
130	AAB28199	Homo sapiens	Human HMG-17 non histone chromosomal protein.	322	75
130	gi306864	Homo sapiens	Human non-histone chromosomal protein HMG-17 mRNA, complete cds.	322	75
130	gi32329	Homo sapiens	Human HMG-17 gene for non-histone chromosomal protein HMG-17.	322	75
131	gi16041794	Homo sapiens	clone MGC:23591 IMAGE:4856946, mRNA, complete cds.	2714	99
131	gi15559462	Homo sapiens	Similar to old astrocyte specifically induced substance, clone MGC:20215 IMAGE:4546950, mRNA, complete cds.	2709	99

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
131	gi4519621	Mus musculus	OASIS protein	2406	91
132	gi7573591	Homo sapiens	Human DNA sequence from clone RP1-309K20 on chromosome 20. Contains the gene for a novel protein similar to dysferlin, the SPAG4 gene for sperm associated antigen 4, the CPNE1 gene for Copine I (similar to KIAA0636), the gene KIAA0765 (HRIHFB2091) for an RNA recognition motif (RNP, RRM or RBD domain) containing protein and the 3' end of the NIFS gene for cysteine desulfurase. Contains ESTs, STSs, GSSs and four putative CpG islands, complete sequence.	4972	100
132	gi15559252	Homo sapiens	RNA binding motif protein 12, clone MGC:19528 IMAGE:3845090, mRNA, complete cds.	4972	100
132	gi15215375	Homo sapiens	RNA binding motif protein 12, clone MGC:16487 IMAGE:3956772, mRNA, complete cds.	4972	100
133	gi12697774	Mus musculus	acetyl-CoA synthetase 2	3181	87
133	gi12697772	Bos taurus	acetyl-CoA synthetase 2	3056	83
133	AAB34712	Homo sapiens	Human secreted protein encoded by DNA clone vo9 1.	2721	100
134	gi7020783	Homo sapiens	cDNA FLJ20580 fis, clone REC00516.	848	100
134	gi15012026	Homo sapiens	Similar to hypothetical protein FLJ20580, clone MGC:13430 IMAGE:4093763, mRNA, complete cds.	848	100
134	gi12833008	Mus musculus	putative	814	85
135	AAB94473	Homo sapiens	Human protein sequence SEQ ID NO:15139.	1970	100
135	AAG74880	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:5644.	1970	100
135	AAB43720	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1165.	1970	100
136	gi10047285	Homo sapiens	mRNA for KIAA1605 protein, partial cds.	3610	99
136	gi16215453	Homo sapiens	mRNA for bile acid beta-glucosidase.	3610	99
136	gi15030210	Homo sapiens	KIAA1605 protein, clone MGC:16895 IMAGE:4339156, mRNA, complete cds.	3610	99
137	gi4914601	Homo sapiens	mRNA; cDNA DKFZp564A026 (from clone DKFZp564A026).	4171	94
137	AAB94357	Homo sapiens	Human protein sequence SEQ ID NO:14881.	2195	99
137	AAY45161	Homo sapiens	Human secreted protein clone CO139 3 protein sequence.	2112	100
138	gi313131	Torpedo marmorata	alpha-tubulin	1192	97
138	gi14198110	Mus musculus	tubulin alpha 1	1192	97
138	gi13435777	Mus musculus	tubulin alpha 6	1192	97

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
139	AAB94856	Homo sapiens	Human protein sequence SEQ ID NO:16044.	2138	100
139	AAB94628	Homo sapiens	Human protein sequence SEQ ID NO:15490.	2138	100
139	gi10436294	Homo sapiens	cDNA FLJ13970 fis, clone Y79AA1001533, moderately similar to Mouse mRNA for RNA polymerase I associated factor (PAF53).	2138	100
140	gi2642187	Rattus norvegicus	endo-alpha-D-mannosidase	1415	67
140	AAB95204	Homo sapiens	Human protein sequence SEQ ID NO:17303.	1094	66
140	gi10434559	Homo sapiens	cDNA FLJ12838 fis, clone NT2RP2003230, moderately similar to Rattus norvegicus endo-alpha-D-mannosidase (Enman) mRNA.	1094	66
141	gi3449308	Homo sapiens	mRNA for MEGF8, partial cds.	9785	100
141	gi6681364	Rattus norvegicus	MEGF8	4772	95
141	gi10728654	Drosophila melanogaster	CG7466 gene product	2902	34
142	AA29517	Homo sapiens	Human lung tumour protein SAL-82 predicted amino acid sequence.	3048	100
142	gi13958036	Homo sapiens	FYVE-finger protein EIP1 mRNA, complete cds.	3048	100
142	AA29861	Homo sapiens	Human secreted protein clone cb98_4.	3041	99
143	gi14718539	Homo sapiens	HIC-3 mRNA, complete cds.	3178	99
143	gi5689371	Homo sapiens	mRNA for KIAA1020 protein, partial cds.	2970	99
143	gi7328028	Homo sapiens	mRNA; cDNA DKFZp434F0616 (from clone DKFZp434F0616); partial cds.	1738	100
144	gi12620400	Homo sapiens	mitochondrial carrier protein CGI-69 long form mRNA, complete cds.	1856	99
144	AAB42783	Homo sapiens	Human ORFX ORF2547 polypeptide sequence SEQ ID NO:5094.	1804	96
144	gi10438783	Homo sapiens	cDNA: FLJ22407 fis, clone HRC08407.	1798	97
145	gi2792366	Homo sapiens	unknown protein IT12 mRNA, partial cds.	4390	99
145	gi1843399	Homo sapiens	mRNA, partial cds, clone:RES4-25.	3676	99
145	gi14602505	Homo sapiens	clone IMAGE:3936655, mRNA, partial cds.	2366	99
146	gi13359167	Homo sapiens	mRNA for KIAA1646 protein, partial cds.	2581	99
146	AA296059	Homo sapiens	Human sphingosine kinase C.	2456	99
146	gi6572330	Homo sapiens	Human DNA sequence from clone 59H18 on chromosome 22. Contains the 3' part of the gene for KIAA0767, a novel gene, ESTs, STSs, GSSs and a putative CpG island, complete sequence.	1627	96
147	gi14043303	Homo sapiens	exonuclease NEF-sp, clone MGC:15944 IMAGE:3537866, mRNA,	4043	100

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			complete cds.		
147	gi13272524	Homo sapiens	exonuclease NEF-sp mRNA, complete cds.	4039	99
147	gi12053043	Homo sapiens	mRNA; cDNA DKFZp434J0315 (from clone DKFZp434J0315); complete cds.	3843	95
148	gi7243037	Homo sapiens	mRNA for KIAA1328 protein, partial cds.	2894	100
148	gi13874541	Macaca fascicularis	hypothetical protein	2492	93
148	gi1335313	Homo sapiens	Human muscle mRNA for embryonic myosin heavy chain (SMHCE).	129	24
149	AAB42399	Homo sapiens	Human ORFX ORF2163 polypeptide sequence SEQ ID NO:4326.	1362	91
149	AAB42366	Homo sapiens	Human ORFX ORF2130 polypeptide sequence SEQ ID NO:4260.	626	100
149	gi7298594	Drosophila melanogaster	CG10189 gene product	223	35
150	AAB95372	Homo sapiens	Human protein sequence SEQ ID NO:17692.	1538	99
150	gi10435150	Homo sapiens	cDNA FLJ13220 fis, clone NT2RP4002047, moderately similar to GTP-BINDING PROTEIN LEPA.	1538	99
150	gi10437720	Homo sapiens	cDNA: FLJ21595 fis, clone COL07069.	1438	100
151	gi3327080	Homo sapiens	mRNA for KIAA0633 protein, partial cds.	6823	99
151	gi857571	Mus musculus	cordon-bleu gene product	1345	81
151	gi6094680	Homo sapiens	PAC clone RP5-1168M19 from 7p12-q11.21, complete sequence.	1342	100
152	gi15451265	Macaca fascicularis	hypothetical protein	2728	98
152	AAB41597	Homo sapiens	Human ORFX ORF1361 polypeptide sequence SEQ ID NO:2722.	2650	100
152	gi5689443	Homo sapiens	mRNA for KIAA1053 protein, partial cds.	2650	100
153	gi14036062	Homo sapiens	unnamed protein product	1930	100
153	AAG81377	Homo sapiens	Human AFP protein sequence SEQ ID NO:272.	1925	99
153	gi12833112	Mus musculus	putative	1727	88
154	gi12832455	Mus musculus	putative	1220	89
154	gi15080314	Homo sapiens	Similar to RIKEN cDNA 0610010D20 gene, clone MGC:20590 IMAGE:4310241, mRNA, complete cds.	514	100
154	gi6002488	Penicillium chrysogenum	hypothetical protein	338	31
155	gi14017889	Homo sapiens	mRNA for KIAA1836 protein, partial cds.	2511	100
155	AAB94592	Homo sapiens	Human protein sequence SEQ ID NO:15402.	972	50
155	gi10435321	Homo sapiens	cDNA FLJ13337 fis, clone OVARC1001880.	972	50
156	gi14550510	Homo sapiens	pseudouridylate synthase 1, clone	2123	100

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			MGC:2736 IMAGE:2822709, mRNA, complete cds.		
156	gi12804097	Homo sapiens	Similar to pseudouridine synthase 1, clone MGC:11268 IMAGE:3943243, mRNA, complete cds.	2123	100
156	gi4455035	Homo sapiens	pseudouridine synthase 1 (PUS1) mRNA, partial cds.	1927	99
157	AAV58052	Homo sapiens	Human protein kinase H2LAU20 protein sequence.	3198	98
157	gi9652080	Homo sapiens	protein kinase DYRK4 (DYRK4) mRNA, partial cds.	2844	100
157	AAW71685	Homo sapiens	Amino acid sequence of human serine/threonine protein kinase.	1909	97
158	gi7300952	Drosophila melanogaster	BcDNA:LD21504 gene product	971	62
158	gi4972728	Drosophila melanogaster	unknown	971	62
158	AAB97646	Homo sapiens	Ribosomal S3 protein 17.	831	99
159	AAU02201	Homo sapiens	Phosphatase 1 protein-like protein, MEM6.	1514	100
159	gi15551577	Homo sapiens	unnamed protein product	1514	100
159	AAB95633	Homo sapiens	Human protein sequence SEQ ID NO:18363.	1510	99
160	gi12804573	Homo sapiens	Similar to CG11334 gene product, clone MGC:3207 IMAGE:3501899, mRNA, complete cds.	1859	100
160	gi12851419	Mus musculus	putative	1590	86
160	gi7302053	Drosophila melanogaster	CG11334 gene product	1046	59
161	gi1580781	Homo sapiens	Human beige-like protein (BGL) mRNA, partial cds.	9734	99
161	gi10180266	Mus musculus	LBA	9333	86
161	gi10257401	Mus musculus	LBA isoform beta	8920	86
162	gi15082589	Homo sapiens	clone MGC:4408 IMAGE:2906200, mRNA, complete cds.	2065	99
162	gi15638615	Arabidopsis thaliana	HEN1	350	37
162	gi13241746	Arabidopsis thaliana	CORYMBOSA2	350	37
163	gi15291227	Drosophila melanogaster	GH13040p	701	40
163	gi7303780	Drosophila melanogaster	CG12214 gene product	701	40
163	AAB95882	Homo sapiens	Human protein sequence SEQ ID NO:18991.	501	100
164	gi3327170	Homo sapiens	mRNA for KIAA0678 protein, partial cds.	5255	100
164	AAB95304	Homo sapiens	Human protein sequence SEQ ID NO:17542.	4431	99
164	gi14134120	Caenorhabditis elegans	endocytosis protein RME-8	2127	42
165	AAB53427	Homo sapiens	Human colon cancer antigen protein sequence SEQ ID NO:967.	813	96

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
165	gi13905098	Mus musculus	B-cell translocation gene 1, anti-proliferative	813	96
165	gi293306	Mus musculus	B-cell translocation gene-1 protein	813	96
166	gi13365897	Macaca fascicularis	hypothetical protein	2501	97
166	AAV02168	Homo sapiens	A facilitative glucose transporter protein GLUT8.	870	99
166	gi13445575	Homo sapiens	facilitative glucose transporter GLUT10 (SLC2A10) mRNA, complete cds.	835	39
167	gi13365897	Macaca fascicularis	hypothetical protein	2173	97
167	AAV02168	Homo sapiens	A facilitative glucose transporter protein GLUT8.	870	99
167	gi13445575	Homo sapiens	facilitative glucose transporter GLUT10 (SLC2A10) mRNA, complete cds.	678	37
168	gi10047251	Homo sapiens	mRNA for KIAA1588 protein, partial cds.	3292	100
168	gi14424704	Homo sapiens	clone MGC:15071 IMAGE:4110510, mRNA, complete cds.	2315	100
168	gi4567179	Homo sapiens	chromosome 19, BAC 37295 (CIT-B-21A4), complete sequence.	1269	43
169	gi15558943	Homo sapiens	guanylate binding protein 4 mRNA, complete cds.	3134	99
169	gi1174187	Mus musculus	purine nucleotide binding protein	2260	70
169	gi193444	Mus musculus	guanylate binding protein	1986	66
170	gi14585859	Homo sapiens	hypothetical protein SB138	1121	100
170	gi6665778	Mus musculus	cyclin ania-6b	1052	92
170	gi12841169	Mus musculus	putative	1052	92
171	AAB64407	Homo sapiens	Amino acid sequence of human intracellular signalling molecule INTRA39.	3394	100
171	AAB71963	Homo sapiens	Human TGF-beta receptor encoded by cDNA clone HFIHY04.	3394	100
171	gi10438113	Homo sapiens	cDNA: FLJ21908 fis, clone HEP03830.	3385	99
172	gi12652533	Homo sapiens	clone MGC:2637 IMAGE:3505128, mRNA, complete cds.	676	89
172	AAB67453	Homo sapiens	Amino acid sequence of a human chaperone polypeptide.	668	88
172	gi9758421	Arabidopsis thaliana	gene_id:MHF15.7~similar to unknown protein~	199	28
173	AAB97025	Homo sapiens	Human colon carcinoma suppressor gene-related protein.	1773	61
173	gi9857318	Homo sapiens	Asef mRNA for APC-stimulated guanine nucleotide exchange factor, complete cds.	1773	61
173	gi8809845	Homo sapiens	chromosome 2q22 RhoGEF mRNA, complete cds.	1700	61
174	gi12052828	Homo sapiens	mRNA; cDNA DKFZp564N1062 (from clone DKFZp564N1062); complete cds.	1601	99
174	gi12850603	Mus musculus	putative	1062	92

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
174	AAB94655	Homo sapiens	Human protein sequence SEQ ID NO:15568.	671	100
175	gi15080282	Homo sapiens	Similar to putative sialoglycoprotease type 2, clone MGC:20293 IMAGE:4121450, mRNA, complete cds.	1747	99
175	gi11071727	Homo sapiens	mRNA for putative sialoglycoprotease type 2.	1707	92
175	gi12847276	Mus musculus	putative	1541	84
176	AAB36628	Homo sapiens	Human FLEXHT-50 protein sequence SEQ ID NO:50.	527	100
176	AAB94208	Homo sapiens	Human protein sequence SEQ ID NO:14557.	527	100
176	AAG01512	Homo sapiens	Human secreted protein, SEQ ID NO: 5593.	527	100
177	gi15929052	Homo sapiens	Similar to RIKEN cDNA 2810442O16 gene, clone MGC:23197 IMAGE:4861869, mRNA, complete cds.	2084	100
177	gi11493155	Homo sapiens	Human DNA sequence from clone RP5-852M4 on chromosome 20. Contains the gene encoding the HBV associated factor, a novel gene similar to Drosophila CG17883, a putative novel gene, two CpG islands, ESTs, GSSs, and STSs, complete sequence.	1952	100
177	gi12840168	Mus musculus	putative	1938	93
178	AAB87034	Homo sapiens	Human secreted protein TANGO 339, SEQ ID NO:3.	1449	100
178	AAAY76266	Homo sapiens	Human secreted protein encoded by gene 10 fragment.	1449	100
178	AAB87135	Homo sapiens	Human secreted protein TANGO 339 F20Y variant, SEQ ID NO:139.	1446	99
179	gi434763	Homo sapiens	Human mRNA for KIAA0120 gene, complete cds.	1048	100
179	gi14424677	Homo sapiens	transgelin 2, clone MGC:15279 IMAGE:4301018, mRNA, complete cds.	1048	100
179	gi9956026	Homo sapiens	clone CDABP0035 mRNA sequence.	1048	100
180	AAB31677	Homo sapiens	Amino acid sequence of a human protein having a hydrophobic domain.	2803	100
180	AAE03346	Homo sapiens	Human gene 19 encoded secreted protein HCRNF14, SEQ ID NO:120.	2803	100
180	AAE03310	Homo sapiens	Human gene 19 encoded secreted protein HCRNF14, SEQ ID NO:84.	2803	100
181	AAB41910	Homo sapiens	Human ORFX ORF1674 polypeptide sequence SEQ ID NO:3348.	1530	99
181	gi5262467	Homo sapiens	mRNA; cDNA DKFZp564I122 (from clone DKFZp564I122).	1530	99
181	gi12849716	Mus musculus	putative	1259	82
182	gi2072972	Homo sapiens	Human L1 element L1.25 p40 and putative p150 genes, complete cds.	497	53
182	AAB64943	Homo sapiens	Human secreted protein sequence	494	54

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			encoded by gene 7 SEQ ID NO:121.		
182	gi5070622	Homo sapiens	retrotransposon L1 insertion in X-linked retinitis pigmentosa locus, complete sequence.	494	53
183	AAB59191	Homo sapiens	Human NADE.	217	47
183	gi8452894	Homo sapiens	p75NTR-associated cell death executor (NADE) mRNA, complete cds.	217	47
183	gi189379	Homo sapiens	Human unknown protein from clone pHGR74 mRNA, complete cds.	217	47
184	AAB88468	Homo sapiens	Human membrane or secretory protein clone PSEC0263.	4931	97
184	gi14272788	Homo sapiens	unnamed protein product	4931	97
184	gi577301	Homo sapiens	Human mRNA for KIAA0090 gene, partial cds.	4650	99
185	AAG64953	Homo sapiens	Human ATP-dependent helicase protein 68.	3169	100
185	gi12052748	Homo sapiens	mRNA; cDNA DKFZp564B1023 (from clone DKFZp564B1023); complete cds.	2716	100
185	gi12836314	Mus musculus	putative	2655	83
186	gi14017781	Homo sapiens	mRNA for KIAA1782 protein, partial cds.	2834	99
186	gi4062983	Mus musculus	Eos protein	2747	95
186	gi11612390	Homo sapiens	zinc finger transcription factor Eos mRNA, complete cds.	2603	98
187	AAB95721	Homo sapiens	Human protein sequence SEQ ID NO:18592.	2419	100
187	gi10436538	Homo sapiens	cDNA FLJ14153 fis, clone NT2RM1000092, weakly similar to MULTIDRUG RESISTANCE PROTEIN 2.	2419	100
187	gi12248763	Homo sapiens	mRNA for SMAP-4, complete cds.	2323	96
188	gi13278906	Homo sapiens	clone MGC:4440 IMAGE:2959536, mRNA, complete cds.	1040	100
188	gi13278819	Homo sapiens	clone MGC:2776 IMAGE:2959536, mRNA, complete cds.	1040	100
188	AAB95829	Homo sapiens	Human protein sequence SEQ ID NO:18847.	618	79
189	gi14602977	Homo sapiens	Similar to KIAA0789 gene product, clone MGC:16602 IMAGE:4110708, mRNA, complete cds.	3100	99
189	gi3043570	Homo sapiens	mRNA for KIAA0523 protein, partial cds.	2564	100
189	gi14133217	Homo sapiens	mRNA for KIAA0789 protein, partial cds.	1463	49
190	gi9717245	Mus musculus	cytoplasmic dynein heavy chain	5569	98
190	gi294543	Rattus norvegicus	dynein heavy chain	5557	98
190	gi402528	Rattus norvegicus	cytoplasmic dynein heavy chain	5535	98
191	gi13537204	Homo sapiens	mRNA for MAST205, complete cds.	6834	98
191	gi406058	Mus musculus	protein kinase	6343	86
191	gi3882335	Homo sapiens	mRNA for KIAA0807 protein, partial	6300	98

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			cds.		
192	gi12847109	Mus musculus	putative	1356	79
192	gi13623271	Homo sapiens	Similar to RIKEN cDNA 2600005P05 gene, clone MGC:11321 IMAGE:3951804, mRNA, complete cds.	1332	100
192	gi12847837	Mus musculus	putative	1170	76
193	gi38149	Pongo pygmaeus	epsilon-globin	397	100
193	gi903731	Gorilla gorilla	epsilon-globin	397	100
193	gi903707	Pan troglodytes	epsilon-globin	397	100
194	AAB74695	Homo sapiens	Human membrane associated protein MEMAP-1.	1799	100
194	AAE01340	Homo sapiens	Human gene 22 encoded secreted protein fragment, SEQ ID NO:205.	1799	100
194	gi15929183	Homo sapiens	modulator of apoptosis 1, clone MGC:9487 IMAGE:3922055, mRNA, complete cds.	1799	100
195	AAG93260	Homo sapiens	Human protein HP10106.	1769	100
195	gi15029765	Mus musculus	RIKEN cDNA 2810039M17 gene	1650	91
195	gi12849932	Mus musculus	putative	1650	91
196	gi14017843	Homo sapiens	mRNA for KIAA1813 protein, partial cds.	3434	100
196	gi15193290	Homo sapiens	LAPSER1 (LAPSER1) mRNA, complete cds.	3309	100
196	gi8217421	Homo sapiens	Human DNA sequence from clone RP11-108L7 on chromosome 10. contains part of the gene for a novel Insulin-like growth factor binding type protein with Kazal-type serine protease inhibitor domain, the gene for a novel protein similar to rat tricarboxylate carrier, the gene for a novel PDZ (DHR, GLGF) domain protein, the gene for a novel protein similar to KIAA0552, KIAA0341 and Fugu hypothetical protein 2, the gene for a novel protein similar to Plasmodium POM1 and C. elegans F46G11.1, a putative novel gene, the SEMA4G gene for semaphorin 4G and a novel gene. Contains ESTs, STSs, GSSs and seven putative CpG islands, complete sequence.	3264	100
197	gi1458241	Caenorhabditis elegans	Hypothetical protein B0507.2	782	39
197	gi12832510	Mus musculus	putative	490	89
197	AAB54014	Homo sapiens	Human pancreatic cancer antigen protein sequence SEQ ID NO:466.	242	100
198	gi500747	Mus musculus	capping protein beta-subunit, isoform 1	1440	98
198	gi212902	Gallus gallus	actin-capping protein Z beta subunit	1432	98
198	gi12805189	Mus musculus	capping protein (actin filament) muscle	1318	92

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			Z-line, beta		
199	gi14017787	Homo sapiens	mRNA for KIAA1785 protein, partial cds.	3195	100
199	gi13436428	Homo sapiens	Similar to feminization 1 a homolog (C. elegans), clone MGC:4216 IMAGE:2957950, mRNA, complete cds.	2197	64
199	gi12836689	Mus musculus	putative	2164	65
200	gi7959811	Homo sapiens	PRO1167	389	100
200	gi2736345	Caenorhabditis elegans	contains similarity to G-coupled protein receptors	69	33
200	gi7504953	Caenorhabditis elegans	hypothetical protein H22D07.1 - Caenorhabditis elegans >	69	33
201	gi12697975	Homo sapiens	mRNA for KIAA1715 protein, partial cds.	2230	100
201	AAB42461	Homo sapiens	Human ORFX ORF2225 polypeptide sequence SEQ ID NO:4450.	1015	100
201	gi12844031	Mus musculus	putative	567	92
202	gi7296176	Drosophila melanogaster	CG2839 gene product	195	27
202	gi10438900	Homo sapiens	cDNA: FLJ22490 fis, clone HRC10983.	184	97
202	gi5824430	Caenorhabditis elegans	cDNA EST yk501h2.5 comes from this gene~cDNA EST yk523d4.5 comes from this gene~cDNA EST yk553f6.5 comes from this gene~cDNA EST yk595g12.5 comes from this gene~cDNA EST yk606g10.5 comes from this gene~cDNA EST yk652f3.5 comes from this gene	182	21
203	AAM00957	Homo sapiens	Human bone marrow protein, SEQ ID NO: 433.	1725	100
203	gi4151807	Rattus norvegicus	membrane-associated guanylate kinase-interacting protein 2 Maguin-2	1484	62
203	gi4151805	Rattus norvegicus	membrane-associated guanylate kinase-interacting protein 1 Maguin-1	1484	62
204	AAM00844	Homo sapiens	Human bone marrow protein, SEQ ID NO: 207.	1051	98
204	gi4151807	Rattus norvegicus	membrane-associated guanylate kinase-interacting protein 2 Maguin-2	779	69
204	gi4151805	Rattus norvegicus	membrane-associated guanylate kinase-interacting protein 1 Maguin-1	779	69
205	AAM00957	Homo sapiens	Human bone marrow protein, SEQ ID NO: 433.	1576	92
205	gi4151807	Rattus norvegicus	membrane-associated guanylate kinase-interacting protein 2 Maguin-2	1349	57
205	gi4151805	Rattus norvegicus	membrane-associated guanylate kinase-interacting protein 1 Maguin-1	1349	57
206	gi7242969	Homo sapiens	mRNA for KIAA1307 protein, partial cds.	8582	99
206	AAM00860	Homo sapiens	Human bone marrow protein, SEQ ID NO: 223.	4841	98
206	gi4426611	Drosophila	pushover	2137	46

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
		melanogaster			
207	AAB62210	Homo sapiens	Human ABCA2 transporter protein.	9835	99
207	gi13173186	Homo sapiens	ABC transporter ABCA2 (ABCA2) mRNA, complete cds.	9835	99
207	gi9957467	Homo sapiens	ATP-binding cassette sub-family A member 2 (ABCA2) mRNA, complete cds.	9835	99
208	AAB94358	Homo sapiens	Human protein sequence SEQ ID NO:14883.	2268	99
208	gi10434632	Homo sapiens	cDNA FLJ12886 fis, clone NT2RP2004041, weakly similar to SYNAPSINS IA AND IB.	2268	99
208	gi12052738	Homo sapiens	mRNA; cDNA DKFZp564H1322 (from clone DKFZp564H1322); complete cds.	2268	99
209	gi14627122	Homo sapiens	Human DNA sequence from clone RP4-583P15 on chromosome 20. Contains ESTs, STSs, GSSs and ten CpG islands. Contains the TNFRSF6B gene for tumor necrosis factor receptor 6b (decoy), the 3' part of the KIAA1088 gene, the ARFRP1 gene for ADP-ribosylation factor related protein 1, two genes for novel proteins, the gene for a GLUT4 enhancer factor and the gene for a novel zinc finger protein similar to rat RIN ZF and the gene for a novel BTB/POZ domain containing zinc finger protein, complete sequence.	2074	99
209	gi13162677	Homo sapiens	GLUT4 enhancer factor mRNA, complete cds.	2055	98
209	gi12655101	Homo sapiens	clone IMAGE:3140406, mRNA, partial cds.	1766	100
210	gi14279329	Homo sapiens	ubiquitin specific protease (USP28) mRNA, complete cds.	4131	92
210	gi7959297	Homo sapiens	mRNA for KIAA1515 protein, partial cds.	3872	100
210	AAB31552	Homo sapiens	A human ubiquitin specific protease 25 (USP25).	2058	48
211	AAB36579	Homo sapiens	Human FLEXHT-1 protein sequence SEQ ID NO:1.	1829	100
211	AAB94048	Homo sapiens	Human protein sequence SEQ ID NO:14211.	1825	99
211	gi10433984	Homo sapiens	cDNA FLJ12475 fis, clone NT2RM1000962.	1825	99
212	gi15824499	Homo sapiens	GalNAc-4-O-sulfotransferase 1 mRNA, complete cds.	2238	100
212	gi11990885	Homo sapiens	GalNAc4ST mRNA for GalNAc 4-sulfotransferase, complete cds.	2238	100
212	gi15559803	Homo sapiens	carbohydrate (N-acetylgalactosamine 4-O) sulfotransferase 8, clone MGC:20987 IMAGE:4635405, mRNA, complete cds.	2238	100

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
213	AAB43387	Homo sapiens	Human ORFX ORF3151 polypeptide sequence SEQ ID NO:6302.	1056	100
213	gi15292317	Drosophila melanogaster	LD46863p	549	50
213	gi7302029	Drosophila melanogaster	CG12054 gene product	549	50
214	gi12843216	Mus musculus	putative	913	84
214	gi14585867	Homo sapiens	hypothetical protein SB145	297	44
214	gi14388386	Macaca fascicularis	hypothetical protein	295	44
215	gi14133219	Homo sapiens	mRNA for KIAA0833 protein, partial cds.	7195	99
215	gi6580410	Homo sapiens	Human DNA sequence from clone RP3-467L1 on chromosome 1p36.21-36.33. Contains the 3' part of gene KIAA0833, the VAMP3 gene for vesicle-associated membrane protein 3 (cellubrevin), the PER3 gene for period (Drosophila) homolog 3 and the gene for urotensin II. Contains two putative CpG islands, ESTs, STSs and GSSs, complete sequence.	3642	99
215	AAB42729	Homo sapiens	Human ORFX ORF2493 polypeptide sequence SEQ ID NO:4986.	997	54
216	gi7293088	Drosophila melanogaster	CG9213 gene product	811	30
216	gi15810333	Arabidopsis thaliana	unknown protein	713	28
216	gi13324888	Caenorhabditis elegans	Hypothetical protein B0361.2	710	34
217	gi2443331	Xenopus laevis	Nfr1	2421	75
217	AAB34944	Homo sapiens	Human secreted protein sequence encoded by gene 20 SEQ ID NO:148.	1129	91
217	gi15292543	Drosophila melanogaster	SD06560p	911	36
218	gi7243111	Homo sapiens	mRNA for KIAA1365 protein, partial cds.	3855	100
218	gi1657758	Rattus norvegicus	densin-180	3640	93
218	gi8570180	Rattus norvegicus	densin-180 variant D	1250	83
219	gi14017839	Homo sapiens	mRNA for KIAA1811 protein, partial cds.	1726	80
219	gi3217028	Homo sapiens	mRNA for putative serine/threonine protein kinase, partial.	1450	84
219	gi7294217	Drosophila melanogaster	CG6114 gene product	1055	70
220	gi7297674	Drosophila melanogaster	CG13139 gene product	942	75
220	gi12857050	Mus musculus	putative	767	62
220	gi15636900	Gallus gallus	avEna neural variant	139	52
221	gi15489242	Homo sapiens	clone IMAGE:3859726, mRNA,	1001	88

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			partial cds.		
221	gi13543991	Homo sapiens	clone IMAGE:3627860, mRNA, partial cds.	1001	88
221	gi12847182	Mus musculus	putative	328	39
222	gi14133209	Homo sapiens	mRNA for KIAA0654 protein, partial cds.	6089	99
222	gi930343	Homo sapiens	Human LAR-interacting protein 1b mRNA, complete cds.	3559	60
222	gi930341	Homo sapiens	Human LAR-interacting protein 1a mRNA, complete cds.	3503	60
223	gi12620207	Homo sapiens	C1orf25 mRNA, complete cds.	3807	98
223	gi9588430	Homo sapiens	Human DNA sequence from clone GS1-120K12 on chromosome 1q25.3-31.2. Contains the gene for ring finger protein DING or BAP-1, an FTH1 (ferritin, heavy polypeptide 1) pseudogene, the 3' end of the gene for a novel protein similar to archael, yeast and worm N2,N2-dimethylguanosine tRNA methyltransferase, ESTs, STSs, GSSs and two putative CpG islands, complete sequence.	2300	98
223	gi12835704	Mus musculus	putative	1420	88
224	gi14595658	Xenopus laevis	LIM protein prickie	2865	67
224	gi10727796	Drosophila melanogaster	esn gene product	698	42
224	gi6634092	Drosophila melanogaster	LIM-domain protein	698	42
225	gi13375149	Homo sapiens	Human DNA sequence from clone RP5-1118M15 on chromosome 20. Contains part of a gene similar to P14 Bos taurus (P14L), a novel gene, ESTs, STSs, GSSs and a CpG Island, complete sequence.	957	99
225	gi7259265	Mus musculus	contains transmembrane (TM) region	314	50
225	AAV53871	Homo sapiens	A human brain-derived signalling factor polypeptide.	299	45
226	gi12803987	Homo sapiens	clone MGC:4174 IMAGE:3634226, mRNA, complete cds.	743	100
226	gi12805417	Mus musculus	Unknown (protein for MGC:7354)	444	66
226	gi12849498	Mus musculus	putative	235	72
227	AAV91629	Homo sapiens	Human secreted protein sequence encoded by gene 23 SEQ ID NO:302.	1391	87
227	gi7677403	Homo sapiens	F-box protein FBG2 (FBG2) mRNA, complete cds.	1391	87
227	AAV83046	Homo sapiens	F-box protein FBP-6.	1333	82
228	gi15079958	Homo sapiens	chromosome 11 open reading frame 24, clone MGC:19741 IMAGE:3614861, mRNA, complete cds.	2231	99
228	gi11527205	Homo sapiens	DM4E3 (C11orf24) mRNA, complete cds.	2224	99

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
228	AAB18965	Homo sapiens	Amino acid sequence of a human transmembrane protein.	2055	99
229	gi15930199	Homo sapiens	Similar to RIKEN cDNA 4921523I18 gene, clone MGC:9467 IMAGE:3914747, mRNA, complete cds.	1451	99
229	gi13278594	Mus musculus	RIKEN cDNA 4921523I18 gene	1440	97
229	gi12856904	Mus musculus	putative	1440	97
230	gi15680131	Homo sapiens	hypothetical protein FLJ12171, clone MGC:19889 IMAGE:4652087, mRNA, complete cds.	1638	100
230	gi14043242	Homo sapiens	hypothetical protein FLJ12171, clone MGC:15694 IMAGE:3351601, mRNA, complete cds.	1638	100
230	AAB93912	Homo sapiens	Human protein sequence SEQ ID NO:13880.	1634	99
231	AAB56947	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1525.	779	100
231	AAB68408	Homo sapiens	Amino acid sequence of a human NOVI polypeptide.	574	100
231	AAAY81695	Homo sapiens	Human PTN protein sequence.	574	100
232	gi11138034	Homo sapiens	mRNA for KIAA1173 protein, complete cds.	2665	100
232	AAG89259	Homo sapiens	Human secreted protein, SEQ ID NO: 379.	2654	99
232	gi12834372	Mus musculus	putative	2427	90
233	AAB98612	Homo sapiens	Human tumour suppressor gene, TSG16, protein.	1706	55
233	gi11596412	Homo sapiens	GAC-1 (GAC-1) mRNA, complete cds.	893	77
233	gi4240237	Homo sapiens	mRNA for KIAA0874 protein, partial cds.	893	77
234	AAB41108	Homo sapiens	Human ORFX ORF872 polypeptide sequence SEQ ID NO:1744.	4170	99
234	gi6331287	Homo sapiens	mRNA for KIAA1274 protein, partial cds.	3936	99
234	gi1545959	Mus musculus	paladin	3560	80
235	gi9368849	Homo sapiens	mRNA; cDNA DKFZp761G2113 (from clone DKFZp761G2113).	972	99
235	gi7293878	Drosophila melanogaster	CG13379 gene product	274	36
235	gi14532482	Arabidopsis thaliana	AT5g58570/mzn1_20	152	31
236	gi3242242	Mus musculus	hyperpolarization-activated cation channel, HAC2	4309	91
236	gi7407645	Rattus norvegicus	hyperpolarization-activated, cyclic nucleotide-gated potassium channel 1	4306	91
236	gi2708316	Mus musculus	brain cyclic nucleotide gated 1; Bcng-1; brain specific ion channel protein	4301	91
237	AAB13370	Homo sapiens	Human brain-associated protein HBAP-1.	1055	100
237	gi9944291	Homo sapiens	TTYH1 mRNA, complete cds.	1055	100
237	gi9651109	Macaca fascicularis	TTYH1	1032	98

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
238	AAU00476	Homo sapiens	Human INTERCEPT 400 protein.	1428	100
238	AAAY79266	Homo sapiens	Human elongase homologue HS3.	1428	100
238	AAB29648	Homo sapiens	Human membrane-associated protein HUMAP-5.	1428	100
239	AAB84885	Homo sapiens	Human protein, SEQ ID 14.	4029	99
239	AAB84882	Homo sapiens	Human protein, SEQ ID 6.	4029	99
239	gi5262593	Homo sapiens	mRNA; cDNA DKFZp434N093 (from clone DKFZp434N093); partial cds.	3684	99
240	gi13477247	Homo sapiens	Similar to RIKEN cDNA 503140M07 gene, clone MGC:13079 IMAGE:3840918, mRNA, complete cds.	2153	100
240	AAB18987	Homo sapiens	Amino acid sequence of a human transmembrane protein.	2148	99
240	gi7670425	Mus musculus	unnamed protein product	1904	89
241	AAG63222	Homo sapiens	Amino acid sequence of a human lipid metabolism enzyme.	2194	100
241	gi14861069	Mus musculus	phosphatidyl inositol phosphate kinase type II gamma	2120	95
241	gi3387798	Rattus norvegicus	phosphatidylinositol 5-phosphate 4-kinase gamma	2087	95
242	gi7295732	Drosophila melanogaster	ft gene product	2915	39
242	gi157409	Drosophila melanogaster	fat protein	2901	39
242	gi10727403	Drosophila melanogaster	ds gene product	2236	34
243	AAF90315_aa 2	Homo sapiens	Winged helix/zinc finger transcription factor FOXP1 cDNA.	819	98
243	AAB82339	Homo sapiens	Winged helix/zinc finger transcription factor FOXP1.	819	98
243	gi12043714	Homo sapiens	clone pAB195 FOXP1 (FOXP1) mRNA, complete cds.	819	98
244	gi10440073	Homo sapiens	cDNA: FLJ23399 fis, clone HEP18254.	2620	100
244	gi7018524	Homo sapiens	mRNA; cDNA DKFZp762K137 (from clone DKFZp762K137); partial cds.	2524	100
244	gi14133227	Homo sapiens	mRNA for KIAA0970 protein, partial cds.	1367	51
245	AAB94855	Homo sapiens	Human protein sequence SEQ ID NO:16042.	1347	100
245	gi10436290	Homo sapiens	cDNA FLJ13968 fis, clone Y79AA1001493, weakly similar to UBIQUITIN-CONJUGATING ENZYME E2-17 KD 9 (EC 6.3.2.19).	1347	100
245	gi16198439	Homo sapiens	hypothetical protein FLJ13855, clone MGC:16842 IMAGE:3915698, mRNA, complete cds.	1347	100
246	gi6330302	Homo sapiens	mRNA for KIAA1185 protein, partial cds.	2041	100
246	AAG74603	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:5367.	1530	97
246	AAB53321	Homo sapiens	Human colon cancer antigen protein sequence SEQ ID NO:861.	1530	97

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
247	gi535390	Macronuclear Homo sapiens	Human cellular retinol binding protein II (CRBP II) mRNA, complete cds.	715	99
247	gi397352	Mus musculus	mCRBP II	674	91
247	gi12833902	Mus musculus	putative	669	90
248	AAG01285	Homo sapiens	Human secreted protein, SEQ ID NO: 5366.	209	87
248	AAR05562	Homo sapiens	Laminin -binding protein encoded by insert from J9 lambda gt10 phage.	209	87
248	gi1149509	Gallus gallus	37kD Laminin receptor precursor /p40 ribosomal associated protein	209	87
249	gi13162226	Homo sapiens	Human DNA sequence from clone RP4-543J19 on chromosome 20 Contains part of the GNAS1 gene encoding guanine nucleotide binding protein (G protein, alpha stimulating activity polypeptide 1) including neuroendocrine secretory protein 55 (NESP55), the CTSZA gene encoding cathepsin Z, the ATP5E gene encoding ATP synthase (H ⁺ transporting, mitochondrial F1 complex, epsilon subunit), the gene encoding protein HSPC130 (TH1 Drosophila homolog), the gene for tubulin beta 1 class VI (TUBB1), a gene encoding the CGI-107 protein (LOC51012), four CpG islands, ESTs, STSs and GSSs, complete sequence.	1591	100
249	gi11230445	Homo sapiens	TUBB1 gene for human beta tubulin 1, class VI.	1591	100
249	gi212834	Gallus gallus	beta-tubulin	1340	85
250	gi13162226	Homo sapiens	Human DNA sequence from clone RP4-543J19 on chromosome 20 Contains part of the GNAS1 gene encoding guanine nucleotide binding protein (G protein, alpha stimulating activity polypeptide 1) including neuroendocrine secretory protein 55 (NESP55), the CTSZA gene encoding cathepsin Z, the ATP5E gene encoding ATP synthase (H ⁺ transporting, mitochondrial F1 complex, epsilon subunit), the gene encoding protein HSPC130 (TH1 Drosophila homolog), the gene for tubulin beta 1 class VI (TUBB1), a gene encoding the CGI-107 protein (LOC51012), four CpG islands, ESTs, STSs and GSSs, complete sequence.	1986	100
250	gi11230445	Homo sapiens	TUBB1 gene for human beta tubulin 1, class VI.	1986	100
250	gi212834	Gallus gallus	beta-tubulin	1699	85
251	gi559325	Homo sapiens	Human mRNA for ATP synthase alpha	1566	99

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			subunit, complete cds.		
251	gi559317	Homo sapiens	Human gene for ATP synthase alpha subunit, complete cds (exon 1 to 12).	1566	99
251	gi34468	Homo sapiens	H.sapiens mRNA for mitochondrial ATP synthase.	1566	99
252	gi559325	Homo sapiens	Human mRNA for ATP synthase alpha subunit, complete cds.	2192	84
252	gi559317	Homo sapiens	Human gene for ATP synthase alpha subunit, complete cds (exon 1 to 12).	2192	84
252	gi34468	Homo sapiens	H.sapiens mRNA for mitochondrial ATP synthase.	2192	84
253	gi14550508	Homo sapiens	Similar to CG8974 gene product, clone MGC:2460 IMAGE:2964524, mRNA, complete cds.	1051	100
253	gi15928691	Mus musculus	Unknown (protein for MGC:19394)	1036	98
253	gi7293133	Drosophila melanogaster	CG8974 gene product	608	66
254	AAE04880	Homo sapiens	Human protease protein-7 (PRTS-7).	2795	100
254	gi14043577	Homo sapiens	hypothetical protein FLJ12455, clone MGC:13149 IMAGE:4298740, mRNA, complete cds.	2795	100
254	AAB94023	Homo sapiens	Human protein sequence SEQ ID NO:14157.	2781	99
255	gi2501855	Homo sapiens	22 kDa actin-binding protein (SM22) gene, complete cds.	937	95
255	gi2340833	Homo sapiens	DNA for SM22 alpha, complete cds.	937	95
255	gi2335047	Homo sapiens	mRNA for SM22 alpha, complete cds.	937	95
256	gi15080204	Homo sapiens	similar to prokaryotic-type class I peptide chain release factors, clone MGC:20261 IMAGE:3029407, mRNA, complete cds.	1948	99
256	gi6706658	Homo sapiens	Human DNA sequence from clone RP1-101K10 on chromosome 6q25-26. Contains a novel gene, the gene for a novel protein similar to Prokaryotic-type class I peptide chain release factors, the 3' end of gene RGS17 (RGSZ2) for regulator of G-protein signaling 17, ESTs, STSs, GSSs and two putative CpG islands, complete sequence.	1940	99
256	gi15680165	Homo sapiens	similar to prokaryotic-type class I peptide chain release factors, clone MGC:20252 IMAGE:4646472, mRNA, complete cds.	1375	98
257	gi15080204	Homo sapiens	similar to prokaryotic-type class I peptide chain release factors, clone MGC:20261 IMAGE:3029407, mRNA, complete cds.	1706	90
257	gi6706658	Homo sapiens	Human DNA sequence from clone RP1-101K10 on chromosome 6q25-26. Contains a novel gene, the gene for a novel protein similar to Prokaryotic-	1698	89

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			type class I peptide chain release factors, the 3' end of gene RGS17 (RGSZ2) for regulator of G-protein signaling 17, ESTs, STSs, GSSs and two putative CpG islands, complete sequence.		
257	gi15680165	Homo sapiens	similar to prokaryotic-type class I peptide chain release factors, clone MGC:20252 IMAGE:4646472, mRNA, complete cds.	1133	85
258	gi7295482	Drosophila melanogaster	CG4603 gene product	616	41
258	gi12322327	Arabidopsis thaliana	unknown protein	451	46
258	gi9454545	Arabidopsis thaliana	Unknown protein	451	46
259	AAB95307	Homo sapiens	Human protein sequence SEQ ID NO:17548.	5011	100
259	gi14042477	Homo sapiens	cDNA FLJ14740 fis, clone NT2RP3002602, weakly similar to PROBABLE PROTEIN DISULFIDE ISOMERASE ER-60 PRECURSOR (EC 5.3.4.1).	5011	100
259	gi15862252	Homo sapiens	unnamed protein product	5008	99
260	gi15079416	Homo sapiens	secreted modular calcium-binding protein 1, clone MGC:19895 IMAGE:4549051, mRNA, complete cds.	2359	100
260	AAB19394	Homo sapiens	Amino acid sequence of a human secreted protein.	2355	99
260	gi10432431	Homo sapiens	mRNA for secreted modular calcium-binding protein (smoc1 gene).	2343	99
261	gi7020475	Homo sapiens	cDNA FLJ20400 fis, clone KAT00587.	1687	100
261	gi1118097	Caenorhabditis elegans	proline and glycine-rich	268	33
261	AAW49723	Homo sapiens	Protein polymer adhesive substrate PPAS1-F.	261	32
262	gi16197949	Drosophila melanogaster	LD21896p	325	29
262	gi7293303	Drosophila melanogaster	CG9089 gene product	325	29
262	gi3170539	Takifugu rubripes	unknown	291	40
263	AAB42525	Homo sapiens	Human ORFX ORF2289 polypeptide sequence SEQ ID NO:4578.	3570	80
263	gi2887497	Homo sapiens	chromosome 19, overlapping cosmids R28707 and R34001, complete sequence.	3570	80
263	AAB42538	Homo sapiens	Human ORFX ORF2302 polypeptide sequence SEQ ID NO:4604.	2835	99
264	gi14017849	Homo sapiens	mRNA for KIAA1816 protein, partial cds.	1637	99
264	gi8655687	Homo sapiens	mRNA; cDNA DKFZp762E1511	892	100

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			(from clone DKFZp762E1511).		
264	gi6979930	Homo sapiens	Mam1 mRNA, partial cds.	315	30
265	gi12836420	Mus musculus	putative	2511	93
265	gi10437002	Homo sapiens	cDNA: FLJ21013 fis, clone CAE05223.	1859	99
265	AAB58385	Homo sapiens	Lung cancer associated polypeptide sequence SEQ ID 723.	1704	99
266	gi14198321	Mus musculus	ribosomal protein L31	543	92
266	gi57115	Rattus norvegicus	ribosomal protein L31 (AA 1-125)	543	92
266	gi14586963	Mus musculus	M75	543	92
267	gi178424	Homo sapiens	Human apolipoprotein A-II mRNA, complete cds.	478	96
267	gi296634	Homo sapiens	Human gene for apolipoprotein AII.	478	96
267	gi296633	Homo sapiens	Human DNA for apolipoprotein A-II.	478	96
268	AAB47184	Homo sapiens	ACPLX protein sequence.	3571	100
268	gi7321168	Homo sapiens	Human DNA sequence from clone RP5-860F19 on chromosome 20p12.3-13 Contains the gene for KIAA1442 (similar to olfactory neuronal transcription factors (COE1, COE2, COE3, EBF3, OLF1)), RPL19 (60S ribosomal protein L19) and HSPC080 pseudogenes, the gene for metallopeptidase (CPX-1) and a novel gene. Contains ESTs, STSs, GSSs and four CpG islands, complete sequence.	3571	100
268	AAB36174	Homo sapiens	Human APG04 protein.	3567	99
269	gi2314829	Homo sapiens	jerky gene product homolog mRNA, complete cds.	1430	59
269	gi10140857	Mus musculus	jerky	752	33
269	AAG62624	Homo sapiens	Human cell nucleus regulatory protein 56.	598	34
270	gi7959227	Homo sapiens	mRNA for KIAA1483 protein, partial cds.	2231	99
270	gi34192	Homo sapiens	Human KUP mRNA for protein with two zinc fingers.	627	39
270	gi13310782	Mus musculus	myoneurin	315	24
271	AAB93814	Homo sapiens	Human protein sequence SEQ ID NO:13604.	1408	97
271	gi10433080	Homo sapiens	cDNA FLJ11753 fis, clone HEMBA1005583.	1408	97
271	AAB41771	Homo sapiens	Human ORFX ORF1535 polypeptide sequence SEQ ID NO:3070.	821	99
272	gi7959197	Homo sapiens	mRNA for KIAA1468 protein, partial cds.	4603	100
272	gi15080502	Homo sapiens	clone MGC:16944 IMAGE:4339646, mRNA, complete cds.	4317	94
272	gi9755831	Arabidopsis thaliana	putative protein	675	27
273	gi15080502	Homo sapiens	clone MGC:16944 IMAGE:4339646, mRNA, complete cds.	4362	98

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
273	gi7959197	Homo sapiens	mRNA for KIAA1468 protein, partial cds.	4360	96
273	gi9755831	Arabidopsis thaliana	putative protein	704	28
274	AAB92483	Homo sapiens	Human protein sequence SEQ ID NO:10570.	2626	100
274	gi7021875	Homo sapiens	cDNA FLJ10051 fis, clone HEMBA1001281.	2626	100
274	gi12837616	Mus musculus	putative	2065	90
275	gi10716076	Homo sapiens	mRNA for testis-abundant finger protein, complete cds.	2739	100
275	gi14043332	Homo sapiens	Similar to ring finger protein 23, clone MGC:2475 IMAGE:3051389, mRNA, complete cds.	2533	94
275	gi10716078	Mus musculus	testis-abundant finger protein	2497	92
276	AAB44673	Homo sapiens	Human secreted protein sequence encoded by gene 33 SEQ ID NO:138.	1014	96
276	gi1747	Oryctolagus cuniculus	trichohyalin	213	22
276	gi13936996	Human herpesvirus 8	ORF73	203	22
277	AAG74326	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:5090.	1101	100
277	AAB56461	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1039.	778	100
277	gi12842930	Mus musculus	putative	688	90
278	gi1020145	Homo sapiens	Human DNA binding protein (HPF2) mRNA, complete cds.	1528	47
278	gi14456631	Homo sapiens	Human DNA sequence from clone RP1-54B20 on chromosome Xp11.1-11.3. Contains the 5' end of a novel SSX family protein gene, two novel KRAB box containing C2H2 type zinc finger protein genes, a KRAB box protein pseudogene, the gene for a novel protein similar to lysozyme C (1,4-beta-N-acetylmuramidase), the ZNF81 gene for zinc finger protein 81 (HFZ20), ESTs, STSs, GSSs and three CpG islands, complete sequence.	1497	55
278	gi498152	Homo sapiens	Human mRNA for KIAA0065 gene, partial cds.	1495	46
279	gi2914676	Homo sapiens	chromosome 16, cosmid clone 360H6 (LANL), complete sequence.	882	35
279	gi14250678	Homo sapiens	clone MGC:10489 IMAGE:3945548, mRNA, complete cds.	882	35
279	gi2342506	Homo sapiens	mRNA for zinc finger protein FPM315, complete cds.	875	35
280	gi434779	Homo sapiens	Human mRNA for KIAA0112 gene, partial cds.	2072	100
280	gi15278392	Homo sapiens	homolog of yeast ribosome biogenesis regulatory protein RRS1, clone MGC:4831 IMAGE:3603972, mRNA,	1905	100

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			complete cds.		
280	gi12804751	Homo sapiens	Similar to regulator for ribosome resistance homolog (S. cerevisiae), clone MGC:2755 IMAGE:2824034, mRNA, complete cds.	1905	100
281	AAB95761	Homo sapiens	Human protein sequence SEQ ID NO:18686.	789	100
281	AAG81272	Homo sapiens	Human AFP protein sequence SEQ ID NO:62.	789	100
281	gi14035852	Homo sapiens	unnamed protein product	789	100
282	gi15080911	Homo sapiens	neo-poly(A) polymerase mRNA, complete cds.	3797	99
282	gi15384858	Homo sapiens	mRNA for poly(A) polymerase gamma (PAPOLG gene).	3797	99
282	gi13641252	Homo sapiens	SRP RNA 3' adenylyating enzyme/pap2 mRNA, complete cds.	3779	99
283	gi6807698	Homo sapiens	mRNA; cDNA DKFZp434A1014 (from clone DKFZp434A1014); partial cds.	1437	85
283	gi12853788	Mus musculus	putative	408	38
283	gi4468790	Xenopus laevis	speedy protein	154	26
284	gi3327062	Homo sapiens	mRNA for KIAA0624 protein, partial cds.	10179	99
284	gi13702612	Staphylococcus aureus subsp. aureus N315	ORFID:SA2447~hypothetical protein, similar to streptococcal hemagglutinin protein	223	19
284	gi14248429	Staphylococcus aureus subsp. aureus Mu50	hypothetical protein	223	19
285	gi12697941	Homo sapiens	mRNA for KIAA1698 protein, partial cds.	4716	100
285	gi7299794	Drosophila melanogaster	CG9591 gene product	290	31
285	AAR99256	Homo sapiens	Natural killer lytic associated protein.	92	40
286	AAG62395	Homo sapiens	Human zinc finger protein 46.	2375	100
286	gi7576274	Homo sapiens	Human DNA sequence from clone RP11-393J16 on chromosome 10. Contains part of the ZNF33A gene for zinc finger protein 33a (KOX 31), a novel gene for a novel KRAB box containing zinc finger gene, a zinc finger pseudogene, ESTs, STSs, GSSs and two putative CpG islands, complete sequence.	2015	100
286	gi881564	Homo sapiens	Human zinc finger containing protein ZNF157 (ZNF157) mRNA, complete cds.	1339	51
287	gi2822143	Homo sapiens	chromosome 19, cosmid R30217, complete sequence.	1838	53
287	gi9968290	Homo sapiens	mRNA for zinc finger protein (ZNF304 gene).	1735	50
287	gi13543419	Homo sapiens	Similar to zinc finger protein 304,	1735	51

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			clone MGC:4079 IMAGE:3530863, mRNA, complete cds.		
288	gi540469	Homo sapiens	(clone HGT26) T cell receptor gamma-chain mRNA, V region.	399	91
288	gi3047024	Homo sapiens	T-cell receptor gamma V1 gene region.	384	100
288	gi339167	Homo sapiens	Human T-cell receptor rearranged gamma-chain gene V-region (V4) (subgroup I).	384	100
289	AA Y69976	Homo sapiens	DHFR-HM protein.	886	93
289	gi182724	Homo sapiens	Human dihydrofolate reductase gene.	886	93
289	gi182717	Homo sapiens	Human dihydrofolate reductase gene, exon 6 and 3' flank.	886	93
290	AAE01782	Homo sapiens	Human gene 13 encoded secreted protein HDPNW93, SEQ ID NO:103.	4269	99
290	gi10437433	Homo sapiens	cDNA: FLJ21347 fis, clone COL02724.	4127	97
290	AAB74693	Homo sapiens	Human protease and protease inhibitor PPIM-26.	3948	99
291	gi6681662	Mus musculus	ENH3	955	90
291	gi12844277	Mus musculus	putative	800	79
291	AA Y12510	Homo sapiens	Human 5' EST secreted protein SEQ ID NO:541.	648	99
292	AAB47327	Homo sapiens	FCTR4.	2798	98
292	gi15141735	Homo sapiens	unnamed protein product	2798	98
292	gi9663126	Homo sapiens	mRNA for chromosome 12 open reading frame 3 (C12orf3).	214	24
293	gi10440367	Homo sapiens	mRNA for FLJ00018 protein, partial cds.	5938	100
293	gi15488570	Homo sapiens	Similar to hypothetical protein FLJ00018, clone MGC:10073 IMAGE:3896004, mRNA, complete cds.	4736	99
293	gi10438857	Homo sapiens	cDNA: FLJ22458 fis, clone HRC10001.	1570	99
294	AAB08948	Homo sapiens	Human secreted protein sequence encoded by gene 21 SEQ ID NO:105.	1601	99
294	AAB08911	Homo sapiens	Human secreted protein sequence encoded by gene 21 SEQ ID NO:68.	1601	99
294	AAB80238	Homo sapiens	Human PRO238 protein.	641	44
295	AAB18457	Homo sapiens	A human TANGO 216 polypeptide clone.	2106	98
295	AAB18447	Homo sapiens	Amino acid sequence of human TANGO 216 polypeptide.	2106	98
295	gi14017381	Homo sapiens	tumor endothelial marker 8 precursor (TEM8) mRNA, complete cds.	1231	57
296	gi14388342	Macaca fascicularis	hypothetical protein	3833	92
296	gi7243195	Homo sapiens	mRNA for KIAA1407 protein, partial cds.	3817	100
296	gi15451319	Macaca fascicularis	hypothetical protein	2408	91
297	gi7243039	Homo sapiens	mRNA for KIAA1329 protein, partial cds.	4761	100

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
297	gi12007720	Mus musculus	VPS10 domain receptor protein SorCS2	4466	88
297	gi7715916	Mus musculus	SorCSb splice variant of the VPS10 domain receptor SorCS	2177	47
298	AAM00812	Homo sapiens	Human bone marrow protein, SEQ ID NO: 175.	1488	99
298	gi12846045	Mus musculus	putative	1387	65
298	AAM00925	Homo sapiens	Human bone marrow protein, SEQ ID NO: 401.	996	100
299	gi7298852	Drosophila melanogaster	CG10068 gene product	609	43
299	gi8655669	Homo sapiens	mRNA; cDNA DKFZp547C176 (from clone DKFZp547C176).	482	52
299	AAB42048	Homo sapiens	Human ORFX ORF1812 polypeptide sequence SEQ ID NO:3624.	325	46
300	gi14043285	Homo sapiens	Similar to KIAA0808 gene product, clone MGC:15880 IMAGE:3529159, mRNA, complete cds.	1306	97
300	gi7263912	Homo sapiens	Human DNA sequence from clone RP5-1108D11 on chromosome 20q12-13.11 Contains part of the gene for a novel protein similar to C. elegans T22C1.7, part of the gene for a novel HMG (high mobility group) box protein similar to KIAA0737, KIAA0808 and TNRC9 (CAGF9), ESTs, STSs, GSSs and two putative CpG islands, complete sequence.	797	96
300	gi3882337	Homo sapiens	mRNA for KIAA0808 protein, complete cds.	767	55
301	gi15430292	Homo sapiens	muscle alpha-kinase (MAK) mRNA, complete cds.	5445	99
301	gi7243041	Homo sapiens	mRNA for KIAA1330 protein, partial cds.	4933	100
301	gi14331137	Mus musculus	myocytic induction/differentiation originator	3684	72
302	gi14550508	Homo sapiens	Similar to CG8974 gene product, clone MGC:2460 IMAGE:2964524, mRNA, complete cds.	589	100
302	gi15928691	Mus musculus	Unknown (protein for MGC:19394)	574	97
302	gi2564951	Mus musculus	unknown	378	72
303	gi7242955	Homo sapiens	mRNA for KIAA1300 protein, partial cds.	9573	99
303	gi6599162	Homo sapiens	mRNA; cDNA DKFZp434N1272 (from clone DKFZp434N1272); partial cds.	1392	98
303	AAG75083	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:5847.	628	92
304	gi1408209	Homo sapiens	Human endogenous retrovirus HERV-K(HML6) proviral clone HML6.17 putative polymerase and envelope genes, partial cds, and 3'LTR.	398	86
304	gi2801455	Mouse	Pr160	176	48

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
		mammary tumor virus			
304	gi6911288	Exogenous mouse mammary tumor virus	Gag-Pro-Pol	176	48
305	gi14269502	Homo sapiens	unconventional myosin 1G valine form (MYO1G) mRNA, MYO1G-V allele, partial cds.	3269	98
305	gi14269504	Homo sapiens	unconventional myosin 1G methionine form (MYO1G) mRNA, MYO1G-M allele, partial cds.	3266	97
305	gi3724141	Rattus norvegicus	myosin I	3130	57
306	gi2145060	Homo sapiens	TTF-I interacting peptide 20 mRNA, partial cds.	2081	99
306	gi2224593	Homo sapiens	Human mRNA for KIAA0326 gene, partial cds.	648	39
306	gi488555	Homo sapiens	Human zinc finger protein ZNF135 mRNA, complete cds.	590	40
307	gi13183883	Homo sapiens	PD-1-ligand 2 protein (PDL2) mRNA, complete cds.	1417	99
307	gi13569410	Homo sapiens	butyrophilin precursor B7-DC mRNA, complete cds.	1417	99
307	AAE01352	Homo sapiens	Human gene 1 encoded secreted protein HDPPA04, SEQ ID NO:74.	1416	99
308	AAB87436	Homo sapiens	Human gene 22 encoded secreted protein fragment, SEQ ID NO:177.	383	100
308	AAB94868	Homo sapiens	Human protein sequence SEQ ID NO:16072.	383	100
308	gi10436314	Homo sapiens	cDNA FLJ13984 fis, clone Y79AA1001846.	383	100
309	AAAY85025	Homo sapiens	Human Rap2 amino acid sequence.	206	33
309	gi4678734	Homo sapiens	Human gene from PACs 37M17 and 305B16, chromosome X, similar to small G proteins, especially RAP-2A.	206	33
309	AAM00956	Homo sapiens	Human bone marrow protein, SEQ ID NO: 432.	205	32
310	gi36905	Homo sapiens	Human mRNA for T-cell receptor alpha-chain HAP50 V(a)8.2-J(a)M.	590	100
310	gi1223888	synthetic construct	T cell receptor alpha chain	586	100
310	gi2358036	Homo sapiens	T-cell receptor alpha delta locus from bases 250472 to 501670 (section 2 of 5) of the Complete Nucleotide Sequence.	586	100
311	AAE01596	Homo sapiens	Human gene 13 encoded secreted protein HCLCJ15, SEQ ID NO:146.	1066	92
311	AAE04136	Homo sapiens	Human gene 6 encoded secreted protein HCLBW50, SEQ ID NO:123.	1066	92
311	gi31135	Homo sapiens	H.sapiens mRNA for elongation factor 1-beta.	1066	92
312	gi7243137	Homo sapiens	mRNA for KIAA1378 protein, partial	2400	99

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			cds.		
312	gi12314036	Homo sapiens	Human DNA sequence from clone RP3-383J4 on chromosome 1q24.1-24.3 Contains part of a gene encoding a kelch motif containing protein, part of a novel gene encoding a protein similar to Aspartyl-TRNA synthetase, a putative novel gene, a 40S ribosomal protein S27 (RPS27) pseudogene, 2 CpG islands, ESTs, STSs and GSSs, complete sequence.	1184	44
312	gi4650844	Homo sapiens	mRNA for Kelch motif containing protein, complete cds.	1176	44
313	gi7019945	Homo sapiens	cDNA FLJ20079 fis, clone COL03057.	1610	83
313	gi12804721	Homo sapiens	clone MGC:2663 IMAGE:3543910, mRNA, complete cds.	1271	48
313	AAB43912	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1357.	1255	45
314	AAB41414	Homo sapiens	Human ORFX ORF1178 polypeptide sequence SEQ ID NO:2356.	5094	97
314	gi6329897	Homo sapiens	mRNA for KIAA1137 protein, partial cds.	4798	98
314	gi14043759	Homo sapiens	clone IMAGE:4111596, mRNA, partial cds.	3906	98
315	AAB28375	Homo sapiens	Human hyperpolarisation-activated channel HAC3.	3686	99
315	gi7959337	Homo sapiens	mRNA for KIAA1535 protein, partial cds.	3665	99
315	gi3242244	Mus musculus	hyperpolarization-activated cation channel, HAC3	3556	96
316	gi14198399	Mus musculus	RIKEN cDNA 1500034J20 gene	837	93
316	gi12854536	Mus musculus	putative	837	93
316	gi14250857	Homo sapiens	Human DNA sequence from clone RP5-1137O17 on chromosome 11p12-14.2 Contains part of a gene similar to putative mitochondrial inner membrane protease subunit 2, a novel mRNA, ESTs and GSSs, complete sequence.	775	100
317	gi10439850	Homo sapiens	cDNA: FLJ23233 fis, clone CAS00458.	1081	50
317	gi9968290	Homo sapiens	mRNA for zinc finger protein (ZNF304 gene).	1039	48
317	gi14249844	Homo sapiens	Similar to hypothetical protein FLJ23233, clone MGC:14876 IMAGE:3544044, mRNA, complete cds.	1037	47
318	gi11863686	Mus musculus	neurobeachin	3371	96
318	gi11863539	Gallus gallus	neurobeachin	2100	89
318	AAB92596	Homo sapiens	Human protein sequence SEQ ID NO:10843.	1721	100
319	gi12698174	Macaca fascicularis	hypothetical protein	1221	95

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
319	gi10439153	Homo sapiens	cDNA: FLJ22672 fis, clone HSI09265.	1085	99
319	gi7020125	Homo sapiens	cDNA FLJ20190 fis, clone COLF0714.	893	50
320	gi2865219	Homo sapiens	integrin binding protein Del-1 (Dell) mRNA, complete cds.	447	100
320	AAW94685	Homo sapiens	Human Del-1 protein.	438	98
320	AAW10365	Homo sapiens	Human developmentally-regulated endothelial cell locus-1 protein.	438	98
321	AAB27246	Homo sapiens	Human EXMAD-24 SEQ ID NO: 24.	2047	100
321	AAB42385	Homo sapiens	Human ORFX ORF2149 polypeptide sequence SEQ ID NO:4298.	2047	100
321	gi52998	Mus musculus	macrophage mannose receptor precursor	164	31
322	gi12834087	Mus musculus	putative	1456	82
322	gi2463628	Homo sapiens	Human putative monocarboxylate transporter (MCT) mRNA, complete cds.	506	29
322	gi2198807	Gallus gallus	monocarboxylate transporter 3	473	27
323	gi15620909	Homo sapiens	mRNA for KIAA1925 protein, partial cds.	1059	38
323	AAB92496	Homo sapiens	Human protein sequence SEQ ID NO:10598.	1050	36
323	gi7021900	Homo sapiens	cDNA FLJ10065 fis, clone HEMBA1001455.	1050	36
324	gi9651075	Macaca fascicularis	unnamed protein product	3716	95
324	gi15145795	Sus scrofa	basic proline-rich protein	222	26
324	gi5917666	Zea mays	extensin-like protein	195	25
325	gi7529597	Homo sapiens	Human DNA sequence from clone RP3-402N21 on chromosome 6p21.1-21.31. Contains up to three novel genes with MAM and immunoglobulin domains. Contains ESTs, STSs, GSSs and four putative CpG islands, complete sequence.	1474	100
325	gi12836077	Mus musculus	putative	1365	95
325	AAE00586	Homo sapiens	Human nuclear cell adhesion molecule homologue, NCAM d 2 protein.	1303	49
326	gi15278193	Homo sapiens	MAGI-1C beta mRNA, complete cds, alternatively spliced.	1492	100
326	gi2702351	Mus musculus	putative membrane-associated guanylate kinase 1	1112	83
326	gi5817255	Homo sapiens	mRNA; cDNA DKFZp434B203 (from clone DKFZp434B203); partial cds.	739	100
327	AAB01432	Homo sapiens	Human TANGO 239 (form 2).	3675	99
327	AAB01426	Homo sapiens	Human TANGO 239.	2700	100
327	AAB00036	Homo sapiens	Human TANGO 239 partial sequence.	2483	97
328	gi7243117	Homo sapiens	mRNA for KIAA1368 protein, partial cds.	5542	100
328	AAV71460	Homo sapiens	Human semaphorin 6A-1.	5422	98
328	gi10187891	Homo sapiens	unnamed protein product	5422	98
329	gi13676461	Macaca fascicularis	hypothetical protein	2193	75

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
329	gi4589566	Homo sapiens	mRNA for KIAA0961 protein, complete cds.	2190	75
329	gi456269	Mus musculus domesticus	zinc finger protein 30	2073	71
330	AAB94295	Homo sapiens	Human protein sequence SEQ ID NO:14747.	3062	99
330	gi10434454	Homo sapiens	cDNA FLJ12768 fis, clone NT2RP2001576, weakly similar to HYPOTHETICAL 62.2 KD PROTEIN C4G8.12C IN CHROMOSOME I.	3062	99
330	gi7291781	Drosophila melanogaster	CG3419 gene product	471	32
331	gi12852801	Mus musculus	putative	1185	95
331	gi12314230	Homo sapiens	Human DNA sequence from clone RP5-846F13 on chromosome 1p21.1-22.1 Contains part of the PPAP2C (phosphatidic acid phosphatase type 2c) gene, ESTs, STSs and GSSs, complete sequence.	975	100
331	gi7020303	Homo sapiens	cDNA FLJ20300 fis, clone HEP06465.	748	56
332	gi12309630	Homo sapiens	Human DNA sequence from clone RP11-438B23 on chromosome 9 Contains a novel gene for a neuronal leucine-rich repeat protein, ESTs, STSs and GSSs, complete sequence.	3138	100
332	AAB31161	Homo sapiens	Amino acid sequence of a human TOLL protein.	2600	86
332	gi13444976	Homo sapiens	unnamed protein product	2600	86
333	gi4240145	Homo sapiens	mRNA for KIAA0828 protein, partial cds.	3226	99
333	gi14249936	Homo sapiens	Similar to S-adenosylhomocysteine hydrolase-like 1, clone IMAGE:3536052, mRNA, partial cds.	3202	100
333	AAW56097	Homo sapiens	Amino acid sequence of the ODD4b5.3 enzyme.	2466	84
334	gi13625385	Homo sapiens	EPI64 (EPI64) mRNA, complete cds.	1026	46
334	AAB95321	Homo sapiens	Human protein sequence SEQ ID NO:17577.	1023	50
334	gi10435007	Homo sapiens	cDNA FLJ13130 fis, clone NT2RP3002972, weakly similar to Halocynthia roretzi mRNA for HrPET-1.	1023	50
335	gi15862408	Homo sapiens	unnamed protein product	2255	95
335	gi13272520	Mus musculus	pancreatitis-induced protein 49	2021	85
335	AAE02778	Homo sapiens	Human PRO-C-MG.64 protein encoded by DNA-C-MG.64-1776 cDNA clone.	1784	95
336	gi15862408	Homo sapiens	unnamed protein product	2281	99
336	gi13272520	Mus musculus	pancreatitis-induced protein 49	2047	88
336	AAE02778	Homo sapiens	Human PRO-C-MG.64 protein encoded by DNA-C-MG.64-1776 cDNA clone.	1810	99
337	gi4545313	Mus musculus	prominin-like protein	1021	77
337	gi15042603	Rattus norvegicus	prominin	647	30

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
337	AAB94028	Homo sapiens	Human protein sequence SEQ ID NO:14170.	642	29
338	gi2978255	Mus musculus	myeloid zinc finger protein-2	212	42
338	AAB54292	Homo sapiens	Human pancreatic cancer antigen protein sequence SEQ ID NO:744.	208	30
338	gi8886436	Homo sapiens	myeloid zinc finger protein 1 splice variants (ZNF42) gene, complete cds, alternatively spliced.	207	42
339	gi3882269	Homo sapiens	mRNA for KIAA0774 protein, partial cds.	5974	99
339	gi12860422	Mus musculus	putative	692	96
339	gi15424451	Homo sapiens	hATIP3	606	36
340	AAB36617	Homo sapiens	Human FLEXHT-39 protein sequence SEQ ID NO:39.	584	100
340	gi8218050	Homo sapiens	Human DNA sequence from clone RP1-187J11 on chromosome 6q11.1-22.33. Contains the gene for a novel protein similar to S. pombe and S. cerevisiae predicted proteins, the gene for a novel protein similar to protein kinase C inhibitors, the 3' end of the gene for a novel protein similar to Drosophila L82 and predicted worm proteins, ESTs, STSs, GSSs and two putative CpG islands, complete sequence.	562	100
340	gi13540300	Mus musculus	nucleolar protein C7B	415	66
341	gi14583268	Homo sapiens	cytoplasmic protein mRNA, complete cds.	628	62
341	gi2104769	Homo sapiens	echinoderm microtubule-associated protein homolog HuEMAP mRNA, complete cds.	560	65
341	gi4406218	Homo sapiens	echinoderm microtubule-associated protein-like EMAP2 mRNA, complete cds.	495	59
342	AAB60099	Homo sapiens	Human transport protein TPPT-19.	1616	93
342	gi7294748	Drosophila melanogaster	CG7616 gene product	580	43
342	gi14714781	Mus musculus	RIKEN cDNA 2610005A10 gene	441	35
343	AAB94374	Homo sapiens	Human protein sequence SEQ ID NO:14915.	3938	99
343	gi10434690	Homo sapiens	cDNA FLJ12921 fis, clone NT2RP2004600.	3938	99
343	gi5689736	Homo sapiens	mRNA for myopodin.	883	34
344	AAAY72604	Homo sapiens	Human Electron Transfer Protein, ETRN-2.	717	100
344	gi10953950	Geochelone carbonaria	alpha-D chain hemoglobin	407	54
344	gi4455876	Cairina moschata	alpha D-globin	398	53
345	AAAY72604	Homo sapiens	Human Electron Transfer Protein, ETRN-2.	668	78
345	gi10953950	Geochelone	alpha-D chain hemoglobin	359	43

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
		carbonaria			
345	gi4455876	Cairina moschata	alpha D-globin	349	41
346	gi8655669	Homo sapiens	mRNA; cDNA DKFZp547C176 (from clone DKFZp547C176).	1053	100
346	AAB42048	Homo sapiens	Human ORFX ORF1812 polypeptide sequence SEQ ID NO:3624.	840	100
346	gi7298852	Drosophila melanogaster	CG10068 gene product	601	40
347	gi15778899	Homo sapiens	Similar to f-box only protein 17, clone MGC:11162 IMAGE:3841901, mRNA, complete cds.	1537	99
347	gi9280060	Macaca fascicularis	unnamed protein product	1435	95
347	gi15214527	Homo sapiens	Similar to f-box only protein 17, clone MGC:9379 IMAGE:3864760, mRNA, complete cds.	857	56
348	AAG64860	Homo sapiens	Heart muscle cell differentiation related protein SEQ ID NO: 61.	1079	90
348	AAB99931	Homo sapiens	Human MesP1 protein sequence SEQ ID NO:61.	1079	90
348	gi13623241	Homo sapiens	Similar to mesoderm posterior 1, clone MGC:10676 IMAGE:3944350, mRNA, complete cds.	1079	90
349	gi4235144	Homo sapiens	chromosome 19, BAC 39498 (CIT-B-26L23), complete sequence.	387	100
349	gi8163824	Homo sapiens	krueppel-like zinc finger protein HZF2 mRNA, complete cds.	290	74
349	AAAY39779	Homo sapiens	CBMACD04 protein sequence.	286	71
350	gi7673618	Mus musculus	ubiquitin specific protease	2016	73
350	gi5689463	Homo sapiens	mRNA for KIAA1063 protein, partial cds.	2000	64
350	gi16198231	Drosophila melanogaster	LD43147p	1188	46
351	gi13540193	Homo sapiens	isopentenyl pyrophosphate isomerase 1 (IDI1), HT009-like protein, and isopentenyl pyrophosphate isomerase type 2 (IDI2) genes, complete cds.	1202	100
351	gi13925766	Homo sapiens	isopentenyl diphosphate dimethylallyl diphosphate isomerase 2 (IDI2) gene, exon 4 and complete cds.	1202	100
351	gi13925769	Homo sapiens	isopentenyl diphosphate dimethylallyl diphosphate isomerase 2 (IDI2) mRNA, complete cds.	1202	100
352	gi13561001	Homo sapiens	Human DNA sequence from clone RP11-528A10 on chromosome 6 Contains an IMPDH1 (IMP (inosine monophosphate) dehydrogenase 1) pseudogene, an RNA helicase pseudogene, a novel gene similar to KIAA0161, ESTs, STSs and GSSs, complete sequence.	950	100
352	gi13991706	Mus musculus	UbcM4-interacting protein 4	655	53

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
352	gi1136384	Homo sapiens	Human mRNA for KIAA0161 gene, complete cds.	651	53
353	gi13561001	Homo sapiens	Human DNA sequence from clone RP11-528A10 on chromosome 6 Contains an IMPDH1 (IMP (inosine monophosphate) dehydrogenase 1) pseudogene, an RNA helicase pseudogene, a novel gene similar to KIAA0161, ESTs, STSs and GSSs, complete sequence.	709	79
353	gi13991706	Mus musculus	UbcM4-interacting protein 4	506	45
353	gi1136384	Homo sapiens	Human mRNA for KIAA0161 gene, complete cds.	502	44
354	AAB74446	Homo sapiens	Human protease-inhibitor like protein.	2759	100
354	gi12053227	Homo sapiens	mRNA; cDNA DKFZp434B044 (from clone DKFZp434B044); complete cds.	2756	99
354	gi15593902	Homo sapiens	unnamed protein product	2743	99
355	AAB94358	Homo sapiens	Human protein sequence SEQ ID NO:14883.	1788	98
355	gi10434632	Homo sapiens	cDNA FLJ12886 fis, clone NT2RP2004041, weakly similar to SYNAPSINS IA AND IB.	1788	98
355	gi12052738	Homo sapiens	mRNA; cDNA DKFZp564H1322 (from clone DKFZp564H1322); complete cds.	1788	98
356	gi13436437	Homo sapiens	Similar to RIKEN cDNA 5730438N18 gene, clone MGC:4399 IMAGE:2905957, mRNA, complete cds.	1634	99
356	gi15030091	Mus musculus	Similar to RIKEN cDNA 5730438N18 gene	1508	91
356	AAB43372	Homo sapiens	Human ORFX ORF3136 polypeptide sequence SEQ ID NO:6272.	1464	91
357	AAB73511	Homo sapiens	Human transferase HTFS-18, SEQ ID NO:18.	1880	99
357	AAG74560	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:5324.	450	98
357	AAG02792	Homo sapiens	Human secreted protein, SEQ ID NO: 6873.	324	96
358	gi7673618	Mus musculus	ubiquitin specific protease	2711	95
358	gi5689463	Homo sapiens	mRNA for KIAA1063 protein, partial cds.	2382	78
358	gi5823525	Drosophila melanogaster	ubiquitin-specific protease nonstop	1305	49
359	AAB94775	Homo sapiens	Human protein sequence SEQ ID NO:15864.	1022	100
359	gi10435984	Homo sapiens	cDNA FLJ13842 fis, clone THYRO1000793.	1022	100
359	gi2340162	Xenopus laevis	dsRBP-ZFa	380	44
360	gi3676086	bacteriophage PS119	gp19	291	59
360	gi1778468	Escherichia	hypothetical protein	287	59

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
		coli			
360	gi1786768	Escherichia coli K12	bacteriophage lambda lysozyme homolog	287	59
361	gi13544003	Homo sapiens	clone IMAGE:3677165, mRNA, partial cds.	2172	88
361	gi3169073	Schizosaccharomyces pombe	phenylalanyl-trna synthetase, mitochondrial precursor	233	33
361	gi13877969	Arabidopsis thaliana	putative phenylalanine-tRNA synthetase	228	35
362	gi293694	Mus musculus	laminin receptor	370	49
362	gi13277921	Mus musculus	laminin receptor 1 (67kD, ribosomal protein SA)	367	49
362	gi4633839	Mus musculus	37kDa oncofetal antigen	367	49
363	gi15082271	Homo sapiens	testes development-related NYD-SP21 mRNA, complete cds.	1876	100
363	gi6807923	Homo sapiens	mRNA; cDNA DKFZp434H092 (from clone DKFZp434H092); partial cds.	1620	100
363	gi7294427	Drosophila melanogaster	CG8797 gene product	118	21
364	AAE01355	Homo sapiens	Human gene 4 encoded secreted protein HRABV43, SEQ ID NO:77.	2724	97
364	gi12836042	Mus musculus	putative	2607	93
364	AAE01380	Homo sapiens	Human gene 4 encoded secreted protein HRABV43, SEQ ID NO:102.	2500	97
365	gi10439688	Homo sapiens	cDNA: FLJ23109 fis, clone LNG07754.	2809	99
365	gi9622093	Mus musculus	E-cadherin binding protein E7	2768	97
365	AAG01765	Homo sapiens	Human secreted protein, SEQ ID NO: 5846.	737	99
366	gi12854995	Mus musculus	putative	844	71
366	gi10241691	Homo sapiens	Novel human gene mapping to chromosome 22.	791	99
366	gi14602790	Homo sapiens	DKFZP566F0546 protein, clone MGC:2444 IMAGE:2822570, mRNA, complete cds.	791	99
367	gi15082283	Homo sapiens	Similar to small glutamine-rich tetratricopeptide repeat (TPR)-containing, clone MGC:10496 IMAGE:3625993, mRNA, complete cds.	720	100
367	gi3377591	Homo sapiens	full length insert cDNA YN88E09.	592	100
367	gi15488015	Homo sapiens	TPR-containing co-chaperone mRNA, complete cds.	450	64
368	gi9104819	Xylella fastidiosa 9a5c	hypothetical protein	151	43
368	AAY59981	Homo sapiens	Human endometrium tumour EST encoded protein 41.	128	46
368	AAE03351	Homo sapiens	Human gene 4 encoded secreted protein fragment, SEQ ID NO:126.	121	58
369	gi5817053	Homo sapiens	mRNA; cDNA DKFZp586D0824 (from clone DKFZp586D0824); partial cds.	571	43
369	gi15530285	Homo sapiens	clone MGC:24275 IMAGE:3950542,	571	43

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			mRNA, complete cds.		
369	gi13569476	Mus musculus	immunity-associated nucleotide 4	540	42
370	gi8453103	Homo sapiens	zinc finger protein mRNA, complete cds.	1296	58
370	gi15012179	Homo sapiens	zinc finger protein 16 (KOX 9), clone MGC:15145 IMAGE:3949487, mRNA, complete cds.	1296	58
370	gi498721	Homo sapiens	H.sapiens HZF10 mRNA for zinc finger protein.	1279	55
371	gi15929964	Homo sapiens	Similar to hypothetical protein FLJ10702, clone MGC:21954 IMAGE:4391821, mRNA, complete cds.	973	100
371	AAB42336	Homo sapiens	Human ORFX ORF2100 polypeptide sequence SEQ ID NO:4200.	932	93
371	AAB93080	Homo sapiens	Human protein sequence SEQ ID NO:11912.	923	91
372	gi7328451	Mus musculus	sialidase	893	44
372	AAB93971	Homo sapiens	Human protein sequence SEQ ID NO:14038.	866	42
372	AAW73964	Homo sapiens	Human sialidase protein sequence.	866	42
373	gi1480005	Mus musculus	Zic4 protein	1490	86
373	AAB14349	Homo sapiens	Human Zic1 protein.	1102	67
373	gi1208429	Homo sapiens	mRNA for Zic protein, complete cds.	1102	67
374	gi12860114	Mus musculus	putative	876	40
374	gi161958	Trypanosoma cruzi	surface antigen	177	23
374	gi1334643	Xenopus laevis	APEG precursor protein	174	26
375	AAV99349	Homo sapiens	Human PRO1110 (UNQ553) amino acid sequence SEQ ID NO:31.	1683	100
375	AAB19729	Homo sapiens	Human SECX Clone 4339264-2 encoded protein.	1683	100
375	AAB15549	Homo sapiens	Human immune system molecule from Incyte clone 2774913.	1683	100
376	gi12746394	Homo sapiens	CUG-BP and ETR-3 like factor 4 (CELF4) mRNA, complete cds.	936	100
376	gi13278792	Homo sapiens	Bruno (Drosophila) -like 4, RNA binding protein, clone MGC:2693 IMAGE:2820541, mRNA, complete cds.	911	98
376	gi12804985	Homo sapiens	Similar to etr1, clone MGC:4320 IMAGE:2820541, mRNA, complete cds.	911	98
377	gi12746394	Homo sapiens	CUG-BP and ETR-3 like factor 4 (CELF4) mRNA, complete cds.	905	89
377	gi13278792	Homo sapiens	Bruno (Drosophila) -like 4, RNA binding protein, clone MGC:2693 IMAGE:2820541, mRNA, complete cds.	880	88
377	gi12804985	Homo sapiens	Similar to etr1, clone MGC:4320 IMAGE:2820541, mRNA, complete cds.	880	88

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
378	gi12841060	Mus musculus	putative	809	75
378	gi7293285	Drosophila melanogaster	CG4768 gene product	239	37
378	gi1938566	Caenorhabditis elegans	Hypothetical protein C48B6.3	123	38
379	gi3880385	Caenorhabditis elegans	predicted using Genefinder~contains similarity to Pfam domain: PF01484 (Nematode cuticle collagen N-terminal domain), Score=51.5, E-value=6.1e-12, N=1~cDNA EST yk94a4.5 comes from this gene~cDNA EST yk94a4.3 comes from this gene~cDNA EST yk68d1.5 comes from this gene~cDNA EST yk68d1.3 comes from this gene	79	35
379	gi6684	Caenorhabditis elegans	unnamed protein product	79	35
379	gi156262	Caenorhabditis elegans	collagen	79	35
380	AAB85365	Homo sapiens	Novel Von Willebrand/thrombosporin-like mature protein sequence.	657	94
380	AAB85364	Homo sapiens	Novel Von Willebrand/thrombosporin-like polypeptide.	657	94
380	gi12836633	Mus musculus	putative	651	59
381	gi15024264	Mus musculus	ribosomal protein L35a	191	53
381	gi57119	Rattus norvegicus	ribosomal protein L35a (aa 1-110)	191	53
381	gi12846322	Mus musculus	putative	191	53
382	gi12835133	Mus musculus	putative	617	71
382	gi7293113	Drosophila melanogaster	CG12379 gene product	283	72
382	gi6042159	Caenorhabditis elegans	Hypothetical protein F53A3.7	226	55
383	AAB81053	Homo sapiens	Human protein HP01640 amino acid sequence.	932	100
383	gi12841896	Mus musculus	putative	925	98
383	gi7303144	Drosophila melanogaster	CG10153 gene product	612	65
384	gi10440373	Homo sapiens	mRNA for FLJ00022 protein, partial cds.	1345	93
384	gi10440396	Homo sapiens	mRNA for FLJ00031 protein, partial cds.	647	88
384	gi1086626	Caenorhabditis elegans	Hypothetical protein C06A6.3	273	33
385	gi12053305	Homo sapiens	mRNA; cDNA DKFZp434G099 (from clone DKFZp434G099); complete cds.	1210	100
385	gi2516239	Mus musculus	Rab33B	1138	94
385	gi12836564	Mus musculus	putative	1138	94
386	gi7243247	Homo sapiens	mRNA for KIAA1433 protein, partial cds.	3232	100
386	AAB94053	Homo sapiens	Human protein sequence SEQ ID NO:14222.	3223	99
386	gi13096872	Mus musculus	Unknown (protein for MGC:7720)	2906	89
387	gi14599491	Homo sapiens	small proline-rich protein 2F (SPRR2F)	458	100

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			gene, complete cds.		
387	gi14599489	Homo sapiens	small proline-rich protein 2E (SPRR2E) gene, complete cds.	444	95
387	gi338423	Homo sapiens	Human small proline rich protein (spr11) mRNA, clone 930.	434	94
388	gi6010699	Rattus norvegicus	F-box protein FBL2	1449	99
388	gi14043139	Homo sapiens	RIKEN cDNA 2610511F20 gene, clone MGC:15482 IMAGE:2987858, mRNA, complete cds.	1383	100
388	gi12848653	Mus musculus	putative	1371	99
389	gi2853265	Rattus norvegicus	jun dimerization protein 2	800	96
389	gi12248392	Mus musculus	transcriptional inhibitory factor	795	95
389	gi6648146	Homo sapiens	chromosome 14 clone CTD-2317F5 map 14q24.3, complete sequence.	481	100
390	gi15277240	Homo sapiens	genomic DNA, chromosome 6p21.3, HLA Class I region, section 17/20.	1296	100
390	gi11875405	Homo sapiens	HZFW1 protein mRNA, complete cds.	1291	99
390	gi11875407	Homo sapiens	HZFW2 protein mRNA, complete cds.	773	99
391	gi6572201	Homo sapiens	Human DNA sequence from clone CITF22-27C3 on chromosome 22q13.1-13.31 Contains a gene for a novel protein (DJ1163J1.2) and part of a gene for a novel protein (DJ1163J1.3, similar to mouse B99), ESTs, STSs and GSSs, complete sequence.	863	100
391	gi4469186	Homo sapiens	Human DNA sequence from clone RP5-1163J1 on chromosome 22q13.2-13.33 Contains the 3' part of a gene for a novel KIAA0279 LIKE EGF-like domain containing protein (similar to mouse Celsr1, rat MEGF2), a novel gene for a protein similar to C. elegans B0035.16 and bacterial tRNA (5-Methylaminomethyl-2-thiouridylate)-Methyltransferases, and the 3' part of a novel gene for a protein similar to mouse B99. Contains ESTs, GSSs and putative CpG islands, complete sequence.	863	100
391	AAB92551	Homo sapiens	Human protein sequence SEQ ID NO:10735.	862	96
392	gi5001720	Mus musculus	odd-skipped related 1 protein	1413	97
392	gi15778246	Mus musculus	odd-skipped related 2	924	66
392	gi15488723	Mus musculus	Unknown (protein for MGC:19171)	924	66
393	AAB94364	Homo sapiens	Human protein sequence SEQ ID NO:14895.	2700	99
393	gi10434650	Homo sapiens	cDNA FLJ12895 fis, clone NT2RP2004187, weakly similar to ZINC FINGER PROTEIN 38.	2700	99
393	gi13623217	Homo sapiens	Similar to hypothetical protein FLJ12895, clone IMAGE:3533093,	2150	99

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			mRNA, partial cds.		
394	gi12053105	Homo sapiens	mRNA; cDNA DKFZp434K111 (from clone DKFZp434K111); complete cds.	3116	100
394	gi2282582	Mus musculus	actin-binding protein	2402	74
394	AAR94386	Homo sapiens	Human neural cell protein marker RR/B.	2400	74
395	gi207145	Rattus norvegicus	synaptotagmin II	2128	95
395	gi7739733	Mus musculus	synaptotagmin II	2121	95
395	gi688412	Mus musculus	synaptotagminII/IP4BP	2121	95
396	gi15487674	Homo sapiens	OSBP-related protein 1 mRNA, complete cds.	3220	99
396	AAB92611	Homo sapiens	Human protein sequence SEQ ID NO:10880.	703	100
396	AAV97291	Homo sapiens	Lipid associated protein (LIPAP) 2764333CD1.	703	100
397	gi11231085	Macaca fascicularis	hypothetical protein	490	76
397	gi2447128	Paramecium bursaria Chlorella virus 1	contains 10 ankyrin-like repeats; similar to human ankyrin, corresponds to Swiss-Prot Accession Number P16157	212	33
397	gi6634025	Homo sapiens	mRNA for KIAA0379 protein, partial cds.	203	38
398	AAB21047	Homo sapiens	Human nucleic acid-binding protein, NuABP-51.	1082	100
398	gi833629	Xenopus laevis	nucleoplasmin	459	49
398	gi64940	Xenopus laevis	nucleoplasmin (AA 1-200)	435	46
399	gi15919272	Homo sapiens	putative forkhead/winged-helix transcription factor (FOXP2) mRNA, complete cds.	596	84
399	gi2565057	Homo sapiens	CAGH44 mRNA, partial cds.	596	84
399	gi14582802	Mus musculus	forkhead-related transcription factor 2	588	82
400	AAB08199	Homo sapiens	Amino acid sequence of human diacylglycerol kinase beta (DAGKbeta).	4217	99
400	gi10279722	Homo sapiens	unnamed protein product	4217	99
400	gi485398	Rattus norvegicus	90kDa-diacylglycerol kinase	4046	95
401	gi7670446	Mus musculus	unnamed protein product	1295	87
401	gi13185203	Homo sapiens	unnamed protein product	799	83
401	AAV31642	Homo sapiens	Human transport-associated protein-4 (TRANP-4).	466	35
402	gi12837990	Mus musculus	putative	985	69
402	gi5668737	Mus musculus	UBE-1c2	661	50
402	AAB94645	Homo sapiens	Human protein sequence SEQ ID NO:15538.	426	52
403	gi10439821	Homo sapiens	cDNA: FLJ23209 fis, clone ADSh00512.	2596	99
403	gi10440353	Homo sapiens	mRNA for FLJ00011 protein, partial	1448	97

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			cds.		
403	gi8217420	Homo sapiens	Human DNA sequence from clone RP11-108L7 on chromosome 10. contains part of the gene for a novel Insulin-like growth factor binding type protein with Kazal-type serine protease inhibitor domain, the gene for a novel protein similar to rat tricarboxylate carrier, the gene for a novel PDZ (DHR, GLGF) domain protein, the gene for a novel protein similar to KIAA0552, KIAA0341 and Fugu hypothetical protein 2, the gene for a novel protein similar to Plasmodium POM1 and C. elegans F46G11.1, a putative novel gene, the SEMA4G gene for semaphorin 4G and a novel gene. Contains ESTs, STSs, GSSs and seven putative CpG islands, complete sequence.	1026	100
404	AAB42219	Homo sapiens	Human ORFX ORF1983 polypeptide sequence SEQ ID NO:3966.	2230	96
404	gi3417297	Homo sapiens	Human Chromosome 16 BAC clone CIT987SK-A-635H12, complete sequence.	2230	96
404	gi15559282	Homo sapiens	clone MGC:20208 IMAGE:3936339, mRNA, complete cds.	1021	53
405	gi13365905	Macaca fascicularis	hypothetical protein	1154	99
405	AAB15537	Homo sapiens	Human immune system molecule from Incyte clone 2751129.	911	100
405	AAB04891	Homo sapiens	Human transporter and ion channel-4 (TRICH-4) protein.	360	39
406	gi262843	Rattus sp.	neurotransmitter transporter	3709	96
406	gi545078	Rattus sp.	Na ⁺ /Cl ⁻ -dependent neurotransmitter transporter	3694	96
406	AAR88390	Homo sapiens	Human neurotransmitter transporter protein.	3668	96
407	AAB31212	Homo sapiens	Amino acid sequence of human polypeptide PRO6004.	728	100
407	AAB44331	Homo sapiens	Human PRO4993 protein sequence SEQ ID NO:612.	717	100
407	gi4519558	Rattus norvegicus	Kilon	667	94
408	gi15277972	Mus musculus	Similar to DnaJ (Hsp40) homolog, subfamily B, member 1	808	49
408	gi7804472	Mus musculus	heat shock protein 40	808	49
408	AAB72675	Homo sapiens	Human HDJ1.	804	48
409	gi12841015	Mus musculus	putative	798	52
409	AAB60114	Homo sapiens	Human transport protein TPPT-34.	787	51
409	gi13435410	Mus musculus	Similar to RIKEN cDNA 1810012H11 gene	768	53
410	gi488555	Homo sapiens	Human zinc finger protein ZNF135	1241	52

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			mRNA, complete cds.		
410	AAAY73346	Homo sapiens	HTRM clone 619699 protein sequence.	1238	49
410	AAB43912	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1357.	1231	49
411	gi837292	Rattus norvegicus	S100A1 gene product	278	59
411	AAB45531	Homo sapiens	Human S100A1 protein.	274	57
411	gi11228039	Homo sapiens	S100A1 cDNA	274	57
412	AAB19851	Homo sapiens	Human muscle-specific protein Ozz.	1504	100
412	gi13929456	Homo sapiens	Human DNA sequence from clone RP3-337O18 on chromosome 20q12-13.1. Contains the PLPT gene encoding Phospholipid Transfer Protein, the PPGB gene coding for Lysosomal Protective Protein precursor (EC 3.4.16.5, Cathepsin A, Carboxypeptidase C) and the gene encoding peroxisomal acyl-CoA thioesterase (PTE1, thioesterase II), four novel genes, the gene for a novel protein similar to Drosophila Neuralized (Neu) and the 5' end of an isoform of the TNNC2 gene for fast troponin C2. Contains three CpG islands, ESTs, STSs and GSSs, complete sequence.	1504	100
412	gi12835750	Mus musculus	putative	1328	89
413	gi12847182	Mus musculus	putative	875	87
413	gi4884173	Homo sapiens	mRNA; cDNA DKFZp564G0982 (from clone DKFZp564G0982); partial cds.	646	100
413	gi10047333	Homo sapiens	mRNA for KIAA1628 protein, partial cds.	346	42
414	gi7959343	Homo sapiens	mRNA for KIAA1538 protein, partial cds.	3286	100
414	AAB42721	Homo sapiens	Human ORFX ORF2485 polypeptide sequence SEQ ID NO:4970.	382	100
414	AAB42764	Homo sapiens	Human ORFX ORF2528 polypeptide sequence SEQ ID NO:5056.	355	41
415	gi14043332	Homo sapiens	Similar to ring finger protein 23, clone MGC:2475 IMAGE:3051389, mRNA, complete cds.	1006	43
415	gi10716078	Mus musculus	testis-abundant finger protein	995	42
415	gi10716076	Homo sapiens	mRNA for testis-abundant finger protein, complete cds.	966	40
416	gi3599509	Mus musculus	rho/rac-interacting citron kinase	1507	61
416	gi3360512	Rattus norvegicus	Citron-K kinase	1505	89
416	gi3599507	Mus musculus	rho/rac-interacting citron kinase short isoform	1503	89
417	gi2358070	Mus musculus	trypsinogen 1	898	65
417	gi603903	Gallus gallus	trypsinogen	408	36
417	gi65163	Xenopus	trypsin precursor	405	38

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
		laevis			
418	gi440127	Rattus norvegicus	cerebroglycan	1132	87
418	AAB44256	Homo sapiens	Human PRO705 (UNQ369) protein sequence SEQ ID NO:109.	570	46
418	AAAY25909	Homo sapiens	Human GPC6 protein.	570	46
419	AAM06489	Homo sapiens	Human foetal protein, SEQ ID NO: 220.	376	82
419	gi12835376	Mus musculus	putative	230	31
419	AAE02058	Homo sapiens	Human four disulfide core domain (FDCD)-containing protein.	222	31
420	AAB42561	Homo sapiens	Human ORFX ORF2325 polypeptide sequence SEQ ID NO:4650.	5075	100
420	gi5419865	Homo sapiens	mRNA; cDNA DKFZp434N074 (from clone DKFZp434N074).	5070	99
420	gi4589532	Homo sapiens	mRNA for KIAA0944 protein, partial cds.	3375	61
421	gi10438804	Homo sapiens	cDNA: FLJ22419 fis, clone HRC08593.	1026	60
421	gi13938187	Homo sapiens	hypothetical protein FLJ22419, clone MGC:14900 IMAGE:3347783, mRNA, complete cds.	1026	60
421	gi6690339	Mus musculus	hematopoietic zinc finger protein	717	47
422	AAB94721	Homo sapiens	Human protein sequence SEQ ID NO:15739.	1678	99
422	gi10435784	Homo sapiens	cDNA FLJ13693 fis, clone PLACE2000111.	1678	99
422	gi5706454	Homo sapiens	mRNA for Natural killer cell p44 related gene 2 (NKp44RG2).	158	29
423	gi15026974	Homo sapiens	mRNA for obscurin (OBSCN gene).	2713	96
423	AAB95162	Homo sapiens	Human protein sequence SEQ ID NO:17205.	1173	86
423	gi13938170	Homo sapiens	clone IMAGE:2961284, mRNA, partial cds.	540	26
424	gi12861364	Mus musculus	putative	523	51
424	AAE02058	Homo sapiens	Human four disulfide core domain (FDCD)-containing protein.	485	38
424	gi12655452	Homo sapiens	mRNA for keratin associated protein 4.7 (KRTAP4.7 gene).	485	40
425	gi12830335	Homo sapiens	Human DNA sequence from clone RP11-55008 on chromosome 20. Contains a novel gene encoding a protein kinase, an RPL7 (60S Ribosomal Protein L7) pseudogene, a CpG island, ESTs, STSs and GSSs, complete sequence.	2062	99
425	AAB65688	Homo sapiens	Novel protein kinase, SEQ ID NO: 216.	1732	100
425	AAB65690	Homo sapiens	Novel protein kinase, SEQ ID NO: 218.	1184	69
426	gi388518	Homo sapiens	Human V beta 5.5 mRNA for a new T cell receptor.	627	95
426	gi36173	Homo sapiens	H.sapiens rearranged T-cell receptor beta chain mRNA.	613	94
426	gi1552509	Homo sapiens	Human germline T-cell receptor beta	606	100

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			chain TCRBV13S1, TCRBV6S8A2T, TCRBV5S6A3N2T, TCRBV13S6A2T, TCRBV6S9P, TCRBV5S3A2T, TCRBV13S8P, TCRBV6S3A1N1T, TCRBV5S2, TCRBV6S6A2T, TCRBV5S7P, TCRBV13S4, TCRBV6S2A1N1T, TCRBV5S4A2T, TCRBV6S4A1, TCRBV23S1A2T, TCRBV12S1A1N2, TCRBV21S2A2, TCRBV8S1, TCRBV8S2A1T, TCRBV8S3, TCRBV16S1A1N1, TCRBV24S1A3T, TCRBV25S1A2PT, TCRBV26S1P, TCRBV18S1, TCRBV17S1A1T, TCRBV2S1, TCRBV10S1P genes from bases 257519 to 472940 (section 2 of 3).		
427	AAE04752	Homo sapiens	Human beta-1,3-galactosyltransferase homologue, ZNSSP8.	434	33
427	gi14597533	Homo sapiens	unnamed protein product	434	33
427	gi14039836	Homo sapiens	beta 1,3 N-acetylglucosaminyltransferase Lc3 synthase mRNA, complete cds.	434	33
428	gi596142	Homo sapiens	Human proteasome subunit LMP7 (allele LMP7C) mRNA, complete cds.	628	49
428	gi38482	Homo sapiens	H.sapiens gene for major histocompatibility complex encoded proteasome subunit LMP7.	624	49
428	gi1054747	Homo sapiens	H.sapiens DMA, DMB, HLA-Z1, IPP2, LMP2, TAP1, LMP7, TAP2, DOB, DQB2 and RING8, 9, 13 and 14 genes.	624	49
429	AAG71415	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1096.	1587	100
429	AAG71594	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1275.	1344	83
429	AAG72476	Homo sapiens	Human OR-like polypeptide query sequence, SEQ ID NO: 2157.	1011	100
430	gi10440063	Homo sapiens	cDNA: FLJ23392 fis, clone HEP17418.	3045	100
430	gi15214571	Mus musculus	Unknown (protein for IMAGE:4207025)	2396	80
430	gi1770528	Homo sapiens	H.sapiens mRNA for translin associated zinc finger protein-1.	687	38
431	gi12859929	Mus musculus	putative	917	96
431	gi15207935	Macaca fascicularis	hypothetical protein	301	96
431	gi1655637	Mus musculus	orf	147	27
432	gi4585414	Bacteriophage 933W	hypothetical protein	408	42
432	gi4499798	Bacteriophage 933W	orf15; homologous to ninG gene	408	42
432	gi5881629	Bacteriophage VT2-Sa	hypothetical protein	408	42
433	gi13161184	Homo sapiens	cytochrome P450 2S1 (CYP2S1) mRNA, complete cds.	2615	100

Table 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
433	AAB93056	Homo sapiens	Human protein sequence SEQ ID NO:11860.	2527	100
433	gi14042396	Homo sapiens	cDNA FLJ14699 fis, clone NT2RP2006571, moderately similar to CYTOCHROME P450 2G1 (EC 1.14.14.1).	2527	100
434	gi13445575	Homo sapiens	facilitative glucose transporter GLUT10 (SLC2A10) mRNA, complete cds.	2752	99
434	gi13603727	Homo sapiens	glucose transporter (GLUT10) mRNA, complete cds.	2752	99
434	gi11065680	Homo sapiens	Novel human gene mapping to chromosome 20, similar to membrane transporters.	2752	99
435	gi13310486	Homo sapiens	C2H2 zinc finger protein (SALL3) gene, complete cds.	6094	99
435	gi6688241	Homo sapiens	SALL3 gene, exons 1a, 2 and 3.	6070	99
435	gi1296845	Mus musculus	spalt protein	5089	84
436	AAG71445	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1126.	1312	85
436	AAG71447	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1128.	924	61
436	gi15293797	Homo sapiens	clone OR6M1 olfactory receptor gene, partial cds.	829	78
437	AAB65297	Homo sapiens	Human PRO9828 protein sequence SEQ ID NO:511.	1360	100
437	AAG89178	Homo sapiens	Human secreted protein, SEQ ID NO: 298.	1360	100
437	AAB84652	Homo sapiens	Amino acid sequence of fibroblast growth factor homologue zFGF12.	1360	100
438	gi53756	Mus musculus	minopontin precursor (AA -66 to 272)	1521	100
438	gi297546	Mus musculus	osteopontin	1516	99
438	gi50864	Mus musculus	T lymphocyte activation protein	1514	99

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
1	PF00204	Zinc finger C-x8-C-x5-C-x3-H type (and similar).	PF00204 11.59 9.700e-12 426-437
1	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 3.667e-09 33-42
2	BL00291	Prion protein.	BL00291A 4.49 8.759e-09 185-220
3	PF01105	emp24/gp25L/p24 family.	PF01105B 25.12 1.000e-40 178-230
4	BL00307	Legume lectins beta-chain proteins.	BL00307G 9.91 8.531e-10 678-689
4	PF00922	Vesiculovirus phosphoprotein.	PF00922A 19.17 8.862e-09 281-315
6	BL01159	WW/rsp5/WWP domain proteins.	BL01159 13.85 6.073e-09 61-76
6	BL00591	Glycosyl hydrolases family 10 proteins.	BL00591G 9.65 9.167e-09 311-323
7	BL01159	WW/rsp5/WWP domain proteins.	BL01159 13.85 6.073e-09 61-76
7	BL00591	Glycosyl hydrolases family 10 proteins.	BL00591G 9.65 9.167e-09 311-323
9	BL00913	Iron-containing alcohol dehydrogenases proteins.	BL00913D 24.20 8.981e-17 170-204 BL00913C 7.62 4.375e-11 136-146 BL00913B 10.94 7.706e-11 86-102
10	BL00913	Iron-containing alcohol dehydrogenases proteins.	BL00913D 24.20 8.981e-17 218-252 BL00913C 7.62 4.375e-11 184-194 BL00913B 10.94 7.706e-11 134-150
11	BL50062	BCL2-like apoptosis inhibitors (spans part of BH3, BH1 and BH).	BL50062C 6.66 8.500e-11 349-358
14	BL01144	Ribosomal protein L31e proteins.	BL01144 25.07 9.069e-26 78-130
15	PF00204	Zinc finger C-x8-C-x5-C-x3-H type (and similar).	PF00204 11.59 6.694e-10 355-366
15	BL00904	Protein prenyltransferases alpha subunit repeat proteins proteins.	BL00904A 8.30 4.000e-09 485-535
15	BL00415	Synapsins proteins.	BL00415N 4.29 6.727e-12 483-527 BL00415N 4.29 2.774e-09 118-600 BL00415P 2.37 4.290e-09 819-855 BL00415Q 2.23 6.534e-09 474-510
15	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 4.500e-14 490-505 PR00049D 0.00 2.500e-12 489-504 PR00049D 0.00 4.000e-12 491-506 PR00049D 0.00 8.201e-11 488-503 PR00049D 0.00 1.205e-10 492-507 PR00049D 0.00 3.746e-09 487-502 PR00049D 0.00 5.271e-09 485-500 PR00049D 0.00 6.644e-09 493-508
15	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 9.022e-13 471-504 DM00215 19.43 1.458e-09 483-516 DM00215 19.43 2.678e-09 469-502 DM00215 19.43 5.424e-09 468-501 DM00215 19.43 8.017e-09 470-503 DM00215 19.43 9.085e-09 466-499 DM00215 19.43 9.237e-09 484-517
15	BL01113	C1q domain proteins.	BL01113A 17.99 9.308e-09 116-143
15	BL00048	Protamine P1 proteins.	BL00048 6.39 5.263e-10 196-223 BL00048 6.39 3.363e-09 262-289 BL00048 6.39 9.112e-09 184-211
17	PR00773	GRPE PROTEIN SIGNATURE	PR00773D 16.14 5.922e-09 215-235

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
23	PD00930	PROTEIN GTPASE DOMAIN ACTIVATION.	PD00930B 33.72 7.300e-26 600-203 PD00930A 25.62 1.514e-16 497-523
23	BL50002	Src homology 3 (SH3) domain proteins profile.	BL50002A 14.19 4.000e-12 727-746
23	PF00182	GTPase-activator protein for Rho-like GTPases	PF00182B 14.20 7.333e-12 549-128
25	BL00375	UDP-glycosyltransferases proteins.	BL00375F 16.99 7.061e-35 291-336 BL00375C 18.27 2.615e-19 126-150 BL00375D 14.56 9.000e-17 192-220 BL00375B 21.22 8.627e-16 67-108 BL00375G 13.01 4.577e-13 390-430
28	BL01170	Ribosomal protein L6e proteins.	BL01170A 12.34 9.143e-40 139-175
28	PD01457	RIBOSOMAL PROTEIN 40S ZINC-FINGER METAL.	PD01457A 16.51 9.845e-09 67-112
29	BL00359	Ribosomal protein L11 proteins.	BL00359B 23.07 4.231e-24 56-97 BL00359C 22.18 6.148e-22 111-145 BL00359A 20.66 4.000e-21 20-56
29	BL01108	Ribosomal protein L24 proteins.	BL01108A 20.33 1.000e-08 40-73
30	PR00983	CYSTEINYL-TRNA SYNTHETASE SIGNATURE	PR00983D 14.16 3.209e-23 270-292 PR00983C 11.27 3.415e-21 239-258 PR00983A 11.10 1.878e-12 75-87
30	BL00178	Aminoacyl-transfer RNA synthetases class-I proteins.	BL00178B 7.11 2.286e-09 314-325
31	PR00718	PHOSPHOLIPASE D SIGNATURE	PR00718E 8.61 1.000e-08 327-351
32	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 6.133e-10 49-58
33	PF00992	Troponin.	PF00992A 16.67 7.972e-10 10-45 PF00992A 16.67 5.145e-09 17-52 PF00992A 16.67 6.684e-09 56-91
34	BL01019	ADP-ribosylation factors family proteins.	BL01019A 13.20 8.000e-11 68-108
34	PR00449	TRANSFORMING PROTEIN P21 RAS SIGNATURE	PR00449C 17.27 4.938e-20 75-98 PR00449A 13.20 1.900e-15 34-56 PR00449E 13.50 6.870e-15 173-196 PR00449B 14.34 1.360e-10 57-74 PR00449D 10.79 5.364e-09 137-151
37	PR00764	COMPLEMENT C9 SIGNATURE	PR00764F 16.89 7.783e-11 204-225
37	DM01077	SEX HORMONE-BINDING GLOBULIN.	DM01077A 16.30 1.165e-10 43-90
37	BL00279	Membrane attack complex components / perforin proteins.	BL00279E 37.11 9.163e-09 187-235
38	PR00832	PAXILLIN SIGNATURE	PR00832B 9.87 6.284e-10 768-792
38	PR00806	VINCULIN SIGNATURE	PR00806A 6.63 9.260e-09 766-777
38	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 8.661e-15 766-781 PR00049D 0.00 3.250e-12 764-779 PR00049D 0.00 7.277e-11 765-780 PR00049D 0.00 8.786e-10 763-778 PR00049D 0.00 9.390e-09 762-777
40	BL00226	Intermediate filaments Proteins.	BL00226D 19.10 3.172e-34 397-444 BL00226B 23.86 5.929e-23 230-278

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
			BL00226C 13.23 4.808e-21 296-327 BL00226A 12.77 5.065e-13 129-144 BL00226B 23.86 6.400e-10 181-229
41	BL00243	Integrins beta chain cysteine-rich domain proteins.	BL00243I 31.77 2.014e-09 156-199 BL00243I 31.77 5.437e-09 159-202 BL00243I 31.77 5.690e-09 30-73
41	BL01208	VWFC domain proteins.	BL01208B 15.83 5.865e-09 184-199
41	BL00203	Vertebrate metallothioneins proteins.	BL00203 13.94 3.670e-11 66-112 BL00203 13.94 4.659e-11 40-86 BL00203 13.94 7.429e-11 70-116 BL00203 13.94 9.505e-11 140-186 BL00203 13.94 2.723e-10 21-67 BL00203 13.94 2.723e-10 61-107 BL00203 13.94 3.147e-10 105-151 BL00203 13.94 4.064e-10 22-68 BL00203 13.94 5.213e-10 161-207 BL00203 13.94 6.457e-10 26-72 BL00203 13.94 7.032e-10 184-230 BL00203 13.94 7.223e-10 80-126 BL00203 13.94 9.043e-10 130-176 BL00203 13.94 1.735e-09 175-221 BL00203 13.94 3.020e-09 150-196 BL00203 13.94 3.204e-09 65-111 BL00203 13.94 3.296e-09 95-141 BL00203 13.94 3.663e-09 135-181 BL00203 13.94 5.041e-09 47-93 BL00203 13.94 5.041e-09 85-131 BL00203 13.94 5.500e-09 100-146 BL00203 13.94 5.867e-09 126-172 BL00203 13.94 5.959e-09 90-136 BL00203 13.94 6.694e-09 170-216 BL00203 13.94 6.878e-09 151-197 BL00203 13.94 6.969e-09 17-63 BL00203 13.94 7.337e-09 115-161 BL00203 13.94 7.429e-09 71-117 BL00203 13.94 7.704e-09 171-217 BL00203 13.94 8.531e-09 155-201 BL00203 13.94 8.714e-09 165-211 BL00203 13.94 9.265e-09 116-162
41	BL00269	Mammalian defensins proteins.	BL00269C 16.52 9.289e-09 28-57 BL00269C 16.52 9.289e-09 72-101
41	PD02283	PROTEIN SPORULATION REPEAT PRECU.	PD02283C 17.54 5.050e-09 138-166 PD02283C 17.54 5.175e-09 24-52 PD02283C 17.54 5.175e-09 68-96 PD02283C 17.54 6.738e-09 113-141 PD02283C 17.54 7.188e-09 163-191 PD02283C 17.54 7.750e-09 173-201 PD02283C 17.54 7.975e-09 128-156 PD02283C 17.54 8.650e-09 148-176 PD02283C 17.54 9.325e-09 118-146
41	BL00799	Granulins proteins.	BL00799D 12.41 7.661e-09 49-96 BL00799G 9.41 1.000e-08 39-80
43	BL00291	Prion protein.	BL00291A 4.49 4.414e-09 47-82
44	PF00084	Sushi domain proteins (SCR repeat proteins.	PF00084B 9.45 7.188e-10 1549-1561
44	BL00142	Neutral zinc metallopeptidases, zinc-	BL00142 8.38 2.286e-09 730-741

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
		binding region proteins.	
44	PR00480	ASTACIN FAMILY SIGNATURE	PR00480B 15.41 3.314e-09 725-744
45	BL00414	Profilin proteins.	BL00414D 15.59 9.182e-10 81-108
48	PR00837	ALLERGEN V5/TPX-1 FAMILY SIGNATURE	PR00837D 11.12 6.023e-09 22-36
48	BL01009	Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 proteins.	BL01009E 13.50 8.204e-09 21-37
49	BL00284	Serpins proteins.	BL00284A 15.64 2.350e-20 85-109 BL00284D 16.34 4.240e-19 323-350 BL00284C 28.56 5.600e-17 216-258 BL00284E 19.15 7.500e-14 408-433 BL00284B 17.99 9.379e-13 189-210
50	BL01283	T-box domain proteins.	BL01283A 24.15 2.125e-39 148-196 BL01283B 23.17 9.438e-34 208-250 BL01283D 11.70 7.868e-31 298-331 BL01283C 13.05 8.448e-16 260-274
50	PR00937	T-BOX DOMAIN SIGNATURE	PR00937A 15.25 9.182e-26 156-181 PR00937D 13.41 7.375e-17 259-274 PR00937B 14.58 8.615e-15 223-237 PR00937E 11.86 8.541e-14 301-315 PR00937F 12.53 1.450e-12 322-331 PR00937C 10.51 1.000e-11 240-250
50	PR00938	BRACHYURY PROTEIN FAMILY SIGNATURE	PR00938C 8.28 6.547e-09 264-282
50	PR00427	INTERLEUKIN-8 RECEPTOR SIGNATURE	PR00427A 16.30 6.776e-09 416-431
51	PD01270	RECEPTOR FC IMMUNOGLOBULIN AFFIN.	PD01270D 24.66 8.054e-09 50-86
52	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 2.543e-13 181-221
52	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245A 18.03 7.682e-11 150-172 PR00245C 7.84 5.286e-10 290-306
52	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237C 15.69 3.700e-09 195-218 PR00237G 19.63 8.535e-09 326-353
53	PR00050	COLD SHOCK PROTEIN SIGNATURE	PR00050A 11.28 3.143e-12 42-58 PR00050C 9.82 9.151e-11 85-104
53	BL00352	'Cold-shock' DNA-binding domain proteins.	BL00352B 23.66 2.881e-13 71-110 BL00352A 12.19 1.327e-10 42-57
56	BL01173	Lipolytic enzymes G-D-X-G family, histidine.	BL01173B 13.27 4.462e-17 140-167 BL01173C 8.98 4.349e-14 182-196 BL01173A 9.41 1.818e-13 454-467 BL01173C 8.98 6.553e-13 495-509 BL01173A 9.41 8.364e-13 107-120
57	PR00321	GAMMA G-PROTEIN (TRANSDUCIN) SIGNATURE	PR00321C 15.39 2.473e-12 123-141
58	PR00937	T-BOX DOMAIN SIGNATURE	PR00937A 15.25 1.000e-24 117-142 PR00937D 13.41 5.500e-18 220-235 PR00937B 14.58 5.235e-13 184-198 PR00937F 12.53 1.450e-12 293-302 PR00937E 11.86 1.918e-12 259-273 PR00937C 10.51 3.133e-11 201-211

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
58	BL01283	T-box domain proteins.	BL01283A 24.15 1.000e-40 109-157 BL01283B 23.17 9.156e-34 169-211 BL01283C 13.05 8.286e-17 221-235 BL01283D 11.70 5.709e-11 269-302
58	PR00938	BRACHYURY PROTEIN FAMILY SIGNATURE	PR00938C 8.28 7.384e-09 225-243
59	PD02059	CORE POLYPROTEIN PROTEIN GAG CONTAINS: P.	PD02059A 28.10 2.694e-09 116-157
63	BL00196	Ribosomal protein L30 proteins.	BL00196 34.38 3.250e-15 46-97
64	BL00226	Intermediate filaments proteins.	BL00226B 23.86 1.205e-31 264-312
64	BL01305	moaA / nifB / pqgE family proteins.	BL01305B 10.95 8.875e-09 78-88
68	DM00892	3 RETROVIRAL PROTEINASE.	DM00892C 23.55 6.727e-13 33-67
69	PR00874	FUNGI-IV METALLOTHIONEIN SIGNATURE	PR00874C 4.37 7.214e-10 68-83
69	PD00866	GLYCOPROTEIN PROTEIN SPIKE E2 PRECURSOR PEPLIMER.	PD00866L 3.73 6.564e-10 1-11 PD00866L 3.73 1.443e-09 26-36
69	BL00026	Chitin recognition or binding domain proteins.	BL00026 12.95 3.013e-09 48-69
69	DM01724	kw ALLERGEN POLLEN CIM1 HOL-LI.	DM01724 8.14 3.250e-09 10-30
69	BL01208	VWFC domain proteins.	BL01208B 15.83 6.838e-09 111-126
69	BL00243	Integrins beta chain cysteine-rich domain proteins.	BL00243I 31.77 4.838e-10 106-149 BL00243I 31.77 7.221e-10 18-61 BL00243I 31.77 1.761e-09 41-84 BL00243I 31.77 3.408e-09 31-74 BL00243I 31.77 7.465e-09 71-114
69	BL00203	Vertebrate metallothioneins proteins.	BL00203 13.94 4.107e-13 66-112 BL00203 13.94 2.138e-12 92-138 BL00203 13.94 1.099e-11 28-74 BL00203 13.94 3.176e-11 82-128 BL00203 13.94 3.374e-11 87-133 BL00203 13.94 5.846e-11 77-123 BL00203 13.94 7.231e-11 102-148 BL00203 13.94 1.670e-10 97-143 BL00203 13.94 2.532e-10 103-149 BL00203 13.94 5.021e-10 88-134 BL00203 13.94 7.128e-10 38-84 BL00203 13.94 7.168e-10 107-153 BL00203 13.94 7.702e-10 73-119 BL00203 13.94 9.426e-10 25-71 BL00203 13.94 1.918e-09 101-147 BL00203 13.94 2.745e-09 27-73 BL00203 13.94 4.031e-09 71-117 BL00203 13.94 4.857e-09 36-82 BL00203 13.94 5.041e-09 98-144 BL00203 13.94 5.154e-09 6-52 BL00203 13.94 6.418e-09 76-122 BL00203 13.94 7.980e-09 91-137 BL00203 13.94 8.255e-09 13-59 BL00203 13.94 8.898e-09 48-94
69	PR00876	NEMATODE METALLOTHIONEIN SIGNATURE	PR00876B 7.66 9.514e-09 80-94
73	PR00875	MOLLUSC METALLOTHIONEIN SIGNATURE	PR00875A 5.83 9.679e-10 17-29

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
74	PR00185	HISTONE H4 SIGNATURE	PR00185B 13.68 8.888e-09 364-384
86	PD00066	PROTEIN ZINC-FINGER METAL-BINDI	PD00066 13.92 7.000e-13 200-213
86	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 6.850e-13 850-867 BL00028 16.07 1.900e-10 184-201 BL00028 16.07 6.100e-10 371-388 BL00028 16.07 6.914e-09 317-334
86	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048B 6.02 7.158e-09 197-207
87	PD02870	RECEPTOR INTERLEUKIN-1 PRECURSOR.	PD02870D 15.74 8.468e-09 358-393
88	BL00048	Protamine P1 proteins.	82 BL00048 6.39 5.500e-10 70-97 BL00048 6.39 2.350e-09 62-89 BL00048 6.39 3.700e-09 60-87 BL00048 6.39 5.050e-09 63-90 BL00048 6.39 6.288e-09 61-88 BL00048 6.39 9.438e-09 71-98
89	PR00320	G-PROTEIN BETA WD-40.REPEAT SIGNATURE	PR00320C 13.01 8.920e-10 202-217 PR00320B 12.19 9.486e-10 202-217 PR00320C 13.01 7.900e-09 292-307 PR00320A 16.74 8.902e-09 202-217
90	BL00453	FKBP-type peptidyl-prolyl cis-trans isomerase proteins.	BL00453B 23.86 3.864e-28 106-140 BL00453A 15.57 1.000e-15 81-96 BL00453C 9.72 1.000e-12 147-160
92	PR00299	ALPHA CRYSTALLIN SIGNATURE	PR00299B 17.53 7.211e-09 324-337
93	PF00676	Dehydrogenase E1 component.	PF00676D 14.40 4.857e-13 421-441 PF00676C 16.88 1.931e-10 389-413 PF00676B 24.71 5.433e-10 192-230
96	BL00824	Elongation factor 1 beta/beta'/delta chain proteins.	BL00824B 9.21 3.919e-09 1472-1492
99	PR00417	PROKARYOTIC DNA TOPOISOMERASE I SIGNATURE	PR00417A 12.66 5.415e-09 866-880
102	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 6.936e-29 17-56
102	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 2.059e-14 435-452 BL00028 16.07 7.353e-14 351-368 BL00028 16.07 2.350e-13 295-312 BL00028 16.07 9.100e-13 491-508 BL00028 16.07 2.174e-12 463-480 BL00028 16.07 8.826e-12 211-228 BL00028 16.07 2.038e-11 379-396 BL00028 16.07 2.385e-11 323-340 BL00028 16.07 3.423e-11 239-256 BL00028 16.07 9.654e-11 407-424 BL00028 16.07 1.000e-10 267-284
102	BL00479	Phorbol esters / diacylglycerol binding domain proteins.	BL00479A 19.86 6.362e-09 366-389
102	PD02462	PROTEIN BOLA TRANSCRIPTION REGULATION AC.	PD02462A 22.48 7.695e-09 204-239
102	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 1.000e-15 460-474 PR00048A 10.52 1.000e-14 432-446 PR00048A 10.52 3.250e-14 320-334

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
			PR00048A 10.52 4.750e-14 348-362 PR00048A 10.52 6.250e-14 376-390 PR00048A 10.52 3.133e-13 292-306 PR00048A 10.52 1.529e-12 488-502 PR00048B 6.02 1.000e-11 336-346 PR00048B 6.02 9.308e-11 224-234 PR00048B 6.02 2.688e-10 476-486 PR00048B 6.02 3.250e-10 280-290 PR00048A 10.52 5.696e-10 404-418 PR00048A 10.52 6.087e-10 264-278 PR00048B 6.02 6.187e-10 420-430 PR00048A 10.52 7.214e-10 236-250 PR00048B 6.02 8.875e-10 364-374 PR00048B 6.02 3.368e-09 171-181 PR00048B 6.02 4.316e-09 448-458
103	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 9.438e-37 10-49
103	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 5.500e-13 413-430 BL00028 16.07 1.000e-12 273-290 BL00028 16.07 1.783e-12 357-374 BL00028 16.07 7.577e-11 301-318 BL00028 16.07 9.308e-11 441-458 BL00028 16.07 9.308e-11 469-486 BL00028 16.07 1.300e-10 329-346
103	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 7.000e-14 354-368 PR00048A 10.52 2.286e-13 298-312 PR00048A 10.52 9.357e-13 270-284 PR00048A 10.52 3.209e-12 410-424 PR00048B 6.02 5.000e-12 286-296 PR00048B 6.02 1.000e-11 342-352 PR00048B 6.02 1.000e-11 370-380 PR00048B 6.02 1.125e-10 314-324 PR00048A 10.52 2.565e-10 466-480 PR00048A 10.52 4.522e-10 438-452 PR00048B 6.02 1.474e-09 454-464 PR00048A 10.52 3.520e-09 326-340 PR00048B 6.02 4.789e-09 482-492
103	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 8.200e-16 289-302 PD00066 13.92 3.769e-15 317-330 PD00066 13.92 6.538e-15 373-386 PD00066 13.92 2.800e-14 345-358 PD00066 13.92 4.600e-14 457-470 PD00066 13.92 4.130e-11 401-414 PD00066 13.92 9.654e-10 429-442 PD00066 13.92 5.200e-09 261-274
103	BL01024	Protein phosphatase 2A regulatory subunit PR55 proteins.	BL01024H 13.88 7.353e-09 163-216
104	PD01781	PROTEASE IMMUNOGLOBULIN PRECURSO.	PD01781B 27.55 8.680e-09 325-369
105	PD01781	PROTEASE IMMUNOGLOBULIN PRECURSO.	PD01781B 27.55 8.680e-09 379-423
107	PR00939	C2HC-TYPE ZINC-FINGER	PR00939B 13.27 3.209e-09 1302-1311

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
		SIGNATURE	
108	PD00066	PROTEIN ZINC-FINGER METAL-BINDL	PD00066 13.92 2.800e-14 279-292 PD00066 13.92 4.600e-14 307-320 PD00066 13.92 1.000e-13 335-348 PD00066 13.92 7.500e-13 363-376
108	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 7.882e-14 319-336 BL00028 16.07 7.300e-13 347-364 BL00028 16.07 4.913e-12 291-308 BL00028 16.07 2.500e-10 263-280 BL00028 16.07 1.257e-09 375-392
108	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 4.214e-13 288-302 PR00048B 6.02 5.000e-12 304-314 PR00048A 10.52 6.824e-12 372-386 PR00048A 10.52 7.353e-12 344-358 PR00048A 10.52 7.158e-11 316-330 PR00048B 6.02 7.231e-11 276-286 PR00048B 6.02 1.000e-09 332-342 PR00048B 6.02 6.211e-09 388-398
108	BL00115	Eukaryotic RNA polymerase II heptapeptide repeat proteins.	BL00115Z 3.12 8.842e-18 96-145 BL00115Z 3.12 7.144e-17 89-138 BL00115Z 3.12 6.888e-16 103-152 BL00115Z 3.12 7.791e-15 82-131 BL00115Z 3.12 3.947e-14 61-110 BL00115Z 3.12 7.292e-14 117-166 BL00115Z 3.12 9.164e-14 110-159 BL00115Z 3.12 1.000e-13 75-124 BL00115Z 3.12 3.871e-13 54-103 BL00115Z 3.12 6.819e-13 68-117 BL00115Z 3.12 4.168e-11 124-173 BL00115Z 3.12 9.651e-10 47-96 BL00115Z 3.12 7.485e-09 71-120 BL00115Z 3.12 9.669e-09 78-127
109	PR00193	MYOSIN HEAVY CHAIN SIGNATURE	PR00193D 14.36 5.680e-33 391-420 PR00193C 12.60 4.789e-32 156-184 PR00193B 11.69 1.692e-26 110-136 PR00193E 19.47 5.500e-21 445-474 PR00193A 15.41 4.130e-20 54-74 PR00193E 19.47 5.091e-12 444-473
110	BL00239	Receptor tyrosine kinase class II proteins.	BL00239B 25.15 2.985e-16 67-115
110	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109B 12.27 8.660e-13 132-151
110	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 4.462e-25 132-163 BL00107B 13.31 6.143e-10 197-213
110	DM00406	GLIADIN.	DM00406 7.73 1.800e-09 818-831
110	BL00904	Protein prenyltransferases alpha subunit repeat proteins proteins.	BL00904A 8.30 5.596e-09 815-865
110	BL00415	Synapsins proteins.	BL00415A 6.15 7.684e-09 796-837
110	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 2.373e-09 801-834 DM00215 19.43 7.712e-09 797-830

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
110	PR00209	ALPHA/BETA GLIADIN FAMILY SIGNATURE	PR00209B 4.88 4.188e-09 817-836 PR00209C 4.56 8.929e-09 790-804
111	BL00678	Trp-Asp (WD) repeat proteins proteins.	BL00678 9.67 2.800e-10 366-377 BL00678 9.67 5.263e-09 417-428 BL00678 9.67 6.211e-09 186-197
111	PR00308	TYPE I ANTIFREEZE PROTEIN SIGNATURE	PR00308C 3.83 8.892e-10 108-118 PR00308C 3.83 8.892e-10 109-119 PR00308C 3.83 8.364e-09 107-117
111	PR00320	G-PROTEIN BETA WD-40 REPEAT SIGNATURE	PR00320A 16.74 4.000e-13 364-379 PR00320B 12.19 7.923e-12 415-430 PR00320A 16.74 5.966e-11 415-430 PR00320C 13.01 7.214e-11 415-430 PR00320C 13.01 9.217e-11 364-379 PR00320A 16.74 9.690e-11 184-199 PR00320B 12.19 3.057e-10 184-199 PR00320C 13.01 6.040e-10 184-199 PR00320B 12.19 6.657e-10 364-379 PR00320B 12.19 1.450e-09 457-472 PR00320C 13.01 2.200e-09 240-255 PR00320A 16.74 4.732e-09 457-472 PR00320A 16.74 6.488e-09 281-296 PR00320C 13.01 1.000e-08 281-296
112	DM00547	1 kw CHROMO BROMODOMAIN SHADOW GLOBAL.	DM00547F 23.43 2.350e-35 384-431 DM00547C 17.30 7.000e-19 23-45 DM00547E 13.94 5.154e-17 135-158 DM00547D 11.60 2.750e-13 105-119
112	BL00315	Dehydrins proteins.	BL00315A 9.35 4.246e-10 1301-1329
112	PF00426	Outer Capsid protein VP4 (Hemagglutinin).	PF00426S 15.67 6.438e-10 1271-1309
112	BL00039	DEAD-box subfamily ATP-dependent helicases proteins.	BL00039D 21.67 6.793e-10 368-414
112	PD02191	I ATP-BINDING NUCLEOSIDE TRANSCR.	PD02191A 13.95 9.036e-10 107-122
112	BL00048	Protamine P1 proteins.	BL00048 6.39 1.900e-09 1257-1284 BL00048 6.39 5.050e-09 1258-1285
112	PF00774	Dihydropyridine sensitive L-type calcium channel (Beta subuni.	PF00774A 16.47 7.130e-09 1280-1326 PF00774A 16.47 7.730e-09 1276-1322
112	BL00115	Eukaryotic RNA polymerase II heptapeptide repeat proteins.	BL00115Z 3.12 3.448e-11 1254-1303 BL00115Z 3.12 3.302e-10 1261-1310 BL00115Z 3.12 4.837e-10 1258-1307 BL00115Z 3.12 7.767e-10 1251-1300 BL00115Z 3.12 8.167e-10 1263-1312 BL00115Z 3.12 8.884e-10 1260-1309 09 1247-1296 BL00115Z 3.12 2.985e-09 1240- 1289 BL00115Z 3.12 5.632e-09 1265-1314 BL00115Z 3.12 8.676e-09 1253-1302 BL00115Z 3.12 9.471e-09 1268-1317 BL00115Z 3.12 9.735e-09 1257-1306
112	PF00186	Flocculin repeat proteins.	PF00186I 9.10 5.290e-13 1279-1309 PF00186I 9.10 6.838e-12 1277-1307

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
			PF00186I 9.10 2.957e-11 1282-1312 PF00186I 9.10 7.496e-11 1276-1306 PF00186I 9.10 5.200e-10 1268-1298 PF00186I 9.10 7.450e-10 1278-1308 PF00186I 9.10 7.450e-10 1280-1310 PF00186I 9.10 4.543e-09 1266-1296 PF00186I 9.10 5.252e-09 1285-1315 PF00186I 9.10 6.031e-09 1272-1302 PF00186I 9.10 6.102e-09 1274-1304 PF00186I 9.10 7.236e-09 1270-1300 PF00186I 9.10 8.016e-09 1261-1291 PF00186I 9.10 9.433e-09 1262-1292 PF00186I 9.10 9.433e-09 1267-1297 PF00186I 9.10 1.000e-08 1256-1286
114	PR00700	PROTEIN TYROSINE PHOSPHATASE SIGNATURE	PR00700D 12.47 8.788e-11 237-256
114	BL00383	Tyrosine specific protein phosphatases proteins.	BL00383E 10.35 5.327e-10 240-251
116	PR00884	RIBOSOMAL PROTEIN HS6 SIGNATURE	PR00884E 8.32 4.750e-09 449-466
117	PD02890	ISOMERASE CHALCONE-FLAVONONE FLAV.	PD02890C 16.14 8.457e-09 200-235
118	BL00226	Intermediate filaments proteins.	BL00226B 23.86 6.513e-10 401-449
118	BL00326	Tropomyosins proteins.	BL00326D 8.76 1.925e-09 196-237
118	BL01160	Kinesin light chain repeat proteins.	BL01160B 19.54 2.678e-09 328-382 BL01160B 19.54 8.932e-09 654-708
119	PD01823	PROTEIN INTERGENIC REGION ABC1 PRECURSOR MITOCHONDRION T.	PD01823C 16.13 7.000e-14 352-373 PD01823B 14.96 3.782e-13 328-348 PD01823D 16.66 6.857e-10 430-451
119	PD01115	PRECURSOR AMPHIBIAN SKIN SIGNAL.	PD01115B 12.92 8.421e-09 268-282
122	BL00854	Proteasome B-type subunits proteins.	BL00854C 29.92 8.435e-19 114-143
124	BL00651	Ribosomal protein L9 proteins.	BL00651A 23.25 8.477e-17 94-134
125	BL01245	RIO1/ZK632.3/MJ0444 family proteins.	BL01245F 18.75 2.373e-23 334-371 BL01245A 14.04 8.342e-23 206-231 BL01245C 13.31 6.564e-15 262-282 BL01245E 15.28 1.000e-12 320-330 BL01245B 11.91 9.809e-10 245-255
128	PR00793	PROLYL AMINOPEPTIDASE (S33) FAMILY SIGNATURE	PR00793C 12.24 1.333e-09 168-183
128	PR00111	ALPHA/BETA HYDROLASE FOLD SIGNATURE	PR00111C 13.46 6.000e-09 182-196
129	BL01160	Kinesin light chain repeat proteins.	BL01160D 10.17 7.077e-09 505-534
129	PD00126	PROTEIN REPEAT DOMAIN TPR NUCLEA.	PD00126A 22.53 1.000e-08 695-716
130	BL00355	HMG14 and HMG17 proteins.	BL00355 5.97 8.412e-32 18-49
130	PR00925	NONHISTONE CHROMOSOMAL PROTEIN HMG17 FAMILY SIGNATURE	PR00925B 3.73 3.400e-16 34-47 PR00925A 5.47 1.750e-15 18-33 PR00925C 5.57 9.824e-09 51-62
131	PR00041	CAMP RESPONSE ELEMENT	PR00041E 7.20 2.976e-13 305-326

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
		BINDING (CREB) PROTEIN SIGNATURE	
131	BL00036	bZIP transcription factors basic domain proteins.	BL00036 9.02 4.103e-09 299-312
132	PR00211	GLUTELIN SIGNATURE	PR00211B 0.86 1.750e-09 205-226 PR00211B 0.86 8.750e-09 199-220
132	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 4.529e-11 201-234 DM00215 19.43 1.804e-10 195-228 DM00215 19.43 2.768e-10 192-225 DM00215 19.43 4.054e-10 202-235 DM00215 19.43 6.304e-10 207-240 DM00215 19.43 7.429e-10 180-213 DM00215 19.43 8.393e-10 196-229 DM00215 19.43 8.714e-10 218-251 DM00215 19.43 6.034e-09 185-218 DM00215 19.43 6.034e-09 219-252 DM00215 19.43 6.492e-09 223-256 DM00215 19.43 7.254e-09 200-233 DM00215 19.43 9.390e-09 189-222 DM00215 19.43 9.695e-09 213-246
133	BL00455	Putative AMP-binding domain proteins.	BL00455 13.31 5.125e-11 293-309
133	PR00154	AMP-BINDING SIGNATURE	PR00154A 8.88 6.276e-09 286-298
136	PD00015	GLYCOPROTEIN PRECURSOR CELL SI.	PD00015A 8.90 6.400e-09 243-251
138	BL00227	Tubulin subunits alpha, beta, and gamma proteins.	BL00227B 19.29 1.000e-40 52-107 BL00227C 25.48 1.000e-40 113-165 BL00227A 24.55 8.200e-36 1-35
140	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 8.377e-13 60-75 PR00049D 0.00 7.500e-10 63-78 PR00049D 0.00 8.071e-10 61-76
140	PR00806	VINCULIN SIGNATURE	PR00806B 4.28 8.440e-09 68-82
140	BL00904	Protein prenyltransferases alpha subunit repeat proteins proteins.	BL00904A 8.30 9.553e-09 60-110
141	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 6.438e-12 1175-1190
141	BL01187	Calcium-binding EGF-like domain proteins pattern proteins.	BL01187B 12.04 5.800e-11 1284-1300 BL01187B 12.04 8.200e-11 180-196
141	BL01248	Laminin-type EGF-like (LE) domain proteins.	BL01248 11.02 4.343e-12 1362-1375 BL01248 11.02 2.350e-11 322-335 BL01248 11.02 4.125e-10 271-284
141	PR00764	COMPLEMENT C9 SIGNATURE	PR00764B 13.56 3.475e-09 1047-1068
141	PR00010	TYPE II EGF-LIKE SIGNATURE	PR00010C 11.16 4.205e-09 185-196
141	BL01113	C1q domain proteins.	BL01113A 17.99 5.673e-09 1621-1210
141	PR00011	TYPE III EGF-LIKE SIGNATURE	PR00011D 14.03 8.895e-12 551-132 PR00011B 13.08 5.846e-11 551-132 PR00011D 14.03 3.215e-10 313-332 PR00011A 14.06 4.214e-10 313-332 PR00011B 13.08 7.783e-10 313-332 PR00011A 14.06 7.781e-09 551-132
141	BL00420	Speract receptor repeat proteins domain	BL00420A 20.42 8.200e-09 1186-1215

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
		proteins.	
141	PD02510	ISOMERASE GALACTOSE-6-PHOSPHATE.	PD02510B 18.31 8.170e-09 548-144
141	PR00261	LOW DENSITY LIPOPROTEIN (LDL) RECEPTOR SIGNATURE	PR00261F 11.57 9.544e-09 1052-1074
141	PR00288	PUROTHIONIN SIGNATURE	PR00288C 10.15 9.165e-09 311-326
142	DM01970	0 kw ZK632.12 YDR313C ENDOSOMAL III.	DM01970B 8.60 4.750e-17 114-565
142	BL01160	Kinesin light chain repeat proteins.	BL01160B 19.54 2.373e-09 203-257
142	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 4.000e-09 559-130
142	BL00422	Granins proteins.	BL00422E 26.86 8.615e-09 462-498
143	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 5.846e-15 141-154 PD00066 13.92 9.217e-11 551-564 PD00066 13.92 6.700e-09 523-536
143	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 9.526e-11 122-136 PR00048A 10.52 2.174e-10 532-546 PR00048A 10.52 6.087e-10 588-164 PR00048B 6.02 7.632e-09 138-148 PR00048A 10.52 8.920e-09 504-518
143	PF00651	BTB (also known as BR-C/Ttk) domain proteins.	PF00651 15.00 8.920e-09 59-72
143	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 7.577e-11 535-114 BL00028 16.07 2.200e-10 125-142 BL00028 16.07 5.800e-10 507-524 BL00028 16.07 8.714e-09 591-170 BL00028 16.07 9.743e-09 444-461
144	PR00926	MITOCHONDRIAL CARRIER PROTEIN SIGNATURE	PR00926F 17.75 3.672e-10 262-285
144	BL00215	Mitochondrial energy transfer proteins.	BL00215A 15.82 7.900e-15 16-41 BL00215A 15.82 8.147e-14 260-285 BL00215A 15.82 1.804e-09 166-191 BL00215B 10.44 5.500e-09 114-127
144	PR00927	ADENINE NUCLEOTIDE TRANSLOCATOR 1 SIGNATURE	PR00927B 14.66 8.644e-09 104-126
147	DM01417	6 kw INDUCING XPMC2 MUSHROOM SPAC22G7.04.	DM01417C 12.93 3.250e-11 267-279 DM01417D 11.08 2.200e-10 306-322
148	BL01160	Kinesin light chain repeat proteins.	BL01160B 19.54 8.378e-10 349-403
151	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 7.807e-11 419-434 PR00049D 0.00 8.125e-11 1284-1299 PR00049D 0.00 3.929e-10 1283-1298 PR00049D 0.00 3.288e-09 417-432
151	BL00904	Protein prenyltransferases alpha subunit repeat proteins proteins.	BL00904A 8.30 3.553e-09 416-466
154	BL00665	Dihydrodipicolinate synthetase proteins.	BL00665D 14.76 1.000e-11 109-132 BL00665C 25.58 5.832e-11 50-101
154	PR00146	DIHYDRODIPICOLINATE SYNTHASE SIGNATURE	PR00146D 16.26 2.525e-10 108-126 PR00146A 12.62 8.615e-09 13-35
156	PD02906	SYNTHASE I PSEUDOURIDYLATE PSEUDOURIDINE LYASE TR.	PD02906C 24.17 9.115e-15 171-206 PD02906B 15.35 4.886e-13 142-155

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
			PD02906D 12.27 1.000e-09 239-249 PD02906A 10.84 8.333e-09 92-105
157	BL00107	Protein kinases ATP-binding region proteins.	BL00107B 13.31 2.286e-11 396-412 BL00107A 18.39 6.148e-11 332-363
157	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109B 12.27 4.938e-09 332-351
160	PF01008	Initiation factor 2 subunit.	PF01008B 25.59 9.171e-36 366-409 PF01008A 20.14 8.676e-12 315-336 PF01008C 12.25 7.382e-10 449-469
161	BL00591	Glycosyl hydrolases family 10 proteins.	BL00591D 8.33 6.167e-09 2099-2112
163	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019B 11.36 7.120e-09 99-113 PR00019B 11.36 7.840e-09 73-87
164	BL00198	Nt-dnaJ domain proteins.	BL00198A 8.07 3.000e-14 143-160
164	PR00187	DNAJ PROTEIN FAMILY SIGNATURE	PR00187A 12.84 8.800e-12 139-159
165	PR00310	ANTI-PROLIFERATIVE PROTEIN BTG1 FAMILY SIGNATURE	PR00310B 10.59 4.000e-39 41-71 PR00310C 12.74 2.256e-33 71-101 PR00310D 9.10 9.820e-33 101-131 PR00310A 11.17 7.000e-27 16-41
165	BL00960	BTG1 family proteins.	BL00960B 24.47 1.000e-40 34-79 BL00960C 12.68 6.745e-21 98-120 BL00960A 10.98 5.304e-12 14-26
166	BL00216	Sugar transport proteins.	BL00216B 27.64 2.688e-21 124-174
166	DM00973	3 kw RESISTANCE BENOMYL YLL028W CYCLOHEXIMIDE.	DM00973A 21.17 4.162e-10 96-133
166	PR00171	SUGAR TRANSPORTER SIGNATURE	PR00171D 12.76 3.520e-13 456-478 PR00171E 14.87 2.750e-09 479-492
166	PR00172	GLUCOSE TRANSPORTER SIGNATURE	PR00172D 9.13 6.513e-09 456-480 BL00216B 27.64 5.198e-20 124-174
167	BL00216	Sugar transport proteins.	
167	DM00973	3 kw RESISTANCE BENOMYL YLL028W CYCLOHEXIMIDE.	DM00973A 21.17 4.162e-10 96-133
168	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 5.929e-32 59-98
168	PD00066	PROTEIN ZINC-FINGER METAL-BINDL	PD00066 13.92 2.385e-15 520-533 PD00066 13.92 2.800e-14 296-309 PD00066 13.92 5.200e-14 240-253 PD00066 13.92 5.200e-14 548-561 PD00066 13.92 9.400e-14 436-449 PD00066 13.92 1.000e-13 324-337 PD00066 13.92 6.143e-12 352-365 PD00066 13.92 6.885e-10 268-281
168	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048B 6.02 6.000e-12 237-247 PR00048A 10.52 6.294e-12 333-347 PR00048A 10.52 6.824e-12 361-375 PR00048A 10.52 9.471e-12 249-263 PR00048A 10.52 4.316e-11 119-133 PR00048A 10.52 4.789e-11 529-543 PR00048A 10.52 6.684e-11 445-459 PR00048A 10.52 8.141e-11 305-319 PR00048B 6.02 6.063e-10 321-331

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
			PR00048B 6.02 6.063e-10 517-527 PR00048A 10.52 7.261e-10 221-235 PR00048B 6.02 7.750e-10 545-117 PR00048B 6.02 1.474e-09 293-303 PR00048A 10.52 2.800e-09 389-403 PR00048A 10.52 1.000e-08 417-431
170	PR00456	RIBOSOMAL PROTEIN P2 SIGNATURE	PR00456E 3.06 2.820e-11 6-21 PR00456E 3.06 7.125e-10 3-18
170	PD02331	CYCLIN CELL CYCLE DIVISION PROTE.	PD02331A 19.76 7.429e-15 93-140 PD02331B 13.43 1.125e-09 174-207
170	PR00833	POLLEN ALLERGEN POA PI SIGNATURE	PR00833H 2.30 5.269e-09 3-18
171	PD00126	PROTEIN REPEAT DOMAIN TPR NUCLEA.	PD00126A 22.53 4.706e-14 140-161 PD00126A 22.53 6.824e-14 289-310
173	BL00741	Guanine-nucleotide dissociation stimulators CDC24 family sign.	BL00741B 14.27 3.418e-11 294-317
173	PR00452	SH3 DOMAIN SIGNATURE	PR00452B 11.65 5.154e-11 86-102
173	PR00497	NEUTROPHIL CYTOSOL FACTOR P40 SIGNATURE	PR00497D 11.91 5.962e-10 91-113
173	PF00564	Octicosapeptide repeat proteins.	PF00564B 24.74 6.442e-09 277-328
175	BL01016	Glycoprotease family proteins.	BL01016C 22.84 5.292e-19 60-105 BL01016H 13.71 6.157e-12 307-317 BL01016E 14.88 3.182e-11 141-169 BL01016D 8.86 6.741e-09 118-131
175	PR00789	O-SIALOGLYCOPROTEIN ENDOPEPTIDASE (M22) METALLO-PROTEASE FAMILY SIGNATURE	PR00789E 12.42 7.128e-14 141-163 PR00789C 16.11 2.707e-12 85-105 PR00789B 10.48 1.205e-09 64-85 PR00789D 8.17 7.151e-09 118-131
176	PR00850	GLYCOSYL HYDROLASE FAMILY 59 SIGNATURE	PR00850B 6.67 5.455e-09 148-173
178	PR00259	TRANSMEMBRANE FOUR FAMILY SIGNATURE	PR00259A 9.27 8.676e-20 17-41 PR00259C 16.40 4.750e-17 85-114 PR00259B 14.81 8.615e-12 58-85 PR00259D 13.50 2.528e-11 235-262
178	BL00421	Transmembrane 4 family proteins.	BL00421B 17.62 6.186e-17 64-103 BL00421A 11.79 6.800e-12 13-32 BL00421E 20.97 1.514e-10 232-262 BL00421C 12.89 3.600e-09 147-159
178	PR00235	HERPESVIRUS MAJOR CAPSID PROTEIN (MCP) SIGNATURE	PR00235A 14.64 8.000e-09 87-111
179	BL01052	Calponin family repeat proteins.	BL01052C 18.51 6.806e-40 87-127 BL01052A 16.12 7.180e-32 3-35 BL01052B 15.31 8.031e-26 52-78 BL01052D 10.26 1.000e-24 174-194
179	PR00890	SMOOTH MUSCLE PROTEIN 22-ALPHA (TRANSGELIN) SIGNATURE	PR00890E 14.34 3.813e-21 135-155 PR00890A 8.61 9.775e-21 34-54 PR00890C 8.22 1.000e-17 84-98 PR00890B 8.75 3.455e-17 62-78 PR00890F 12.92 4.064e-14 161-174 PR00890D 16.17 5.174e-13 118-128

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
179	PR00888	SMOOTH MUSCLE PROTEIN/CALPONIN FAMILY SIGNATURE	PR00888H 9.97 5.154e-20 175-191 PR00888C 12.27 5.179e-18 52-68 PR00888D 16.09 4.273e-17 88-105 PR00888A 11.87 2.350e-16 3-18 PR00888E 11.81 3.432e-16 104-120 PR00888F 7.44 4.825e-14 125-140 PR00888G 12.73 8.759e-14 162-176 PR00888B 13.72 2.350e-12 22-36
179	PR00889	CALPONIN SIGNATURE	PR00889E 12.18 2.726e-12 171-187
180	BL00875	Bacterial type II secretion system protein D proteins.	BL00875A 25.57 6.447e-09 367-399
181	PD01351	PROTEIN REPEAT NEUROFILAMENT TRIPL.	PD01351B 13.72 5.355e-09 238-264
182	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354H 18.00 8.826e-27 109-149 DM01354G 11.57 2.143e-25 78-109 DM01354F 14.56 1.414e-15 42-78 DM01354E 18.69 8.650e-14 17-47
182	BL00869	Renal dipeptidase proteins.	BL00869D 14.02 3.477e-09 67-96
185	BL00039	DEAD-box subfamily ATP-dependent helicases proteins.	BL00039A 18.44 4.000e-25 222-261 BL00039D 21.67 4.529e-23 498-544 BL00039C 15.63 4.300e-16 347-371 BL00039B 19.19 9.379e-15 262-288
185	PD00302	PROTEASE POLYPROTEIN HYDROLASE ASP.	PD00302B 9.52 1.346e-09 234-250
186	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 5.714e-12 152-165 PD00066 13.92 6.143e-12 124-137
186	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 6.885e-11 136-153 BL00028 16.07 2.200e-10 197-214
186	PR00239	MOLLUSCAN RHODOPSIN C-TERMINAL TAIL SIGNATURE	PR00239E 1.58 5.705e-09 420-432
186	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 2.957e-10 133-147 PR00048A 10.52 3.739e-10 194-208 PR00048A 10.52 8.043e-10 161-175 PR00048B 6.02 8.105e-09 121-131
187	BL01022	PTR2 family proton/oligopeptide symporters proteins.	BL01022B 22.19 4.240e-10 308-354
187	PR00669	INHIBIN ALPHA CHAIN SIGNATURE	PR00669B 8.27 7.915e-09 264-281
190	PR00830	ENDOPEPTIDASE LA (LON) SERINE PROTEASE (S16) SIGNATURE	PR00830A 8.41 3.342e-09 881-901
191	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109B 12.27 9.234e-13 261-280
191	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 1.000e-23 261-292 BL00107B 13.31 1.000e-12 341-357
191	BL00239	Receptor tyrosine kinase class II proteins.	BL00239B 25.15 6.523e-10 196-244
191	BL00479	Phorbol esters / diacylglycerol binding domain proteins	BL00479C 12.01 1.000e-09 320-333
191	PR00834	HTRA/DEGQ PROTEASE FAMILY	PR00834F 10.91 2.946e-09 786-799

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
		SIGNATURE	
193	BL01033	Globins profile.	BL01033A 16.94 2.385e-18 25-47
193	PR00814	BETA HAEMOGLOBIN SIGNATURE	PR00814A 12.94 1.000e-22 30-47 PR00814B 9.18 7.750e-18 48-64
193	PR00175	MYOGLOBIN SIGNATURE	PR00175B 9.02 9.392e-10 25-49
194	PR00320	G-PROTEIN BETA WD-40 REPEAT SIGNATURE	PR00320B 12.19 6.226e-11 140-155 PR00320A 16.74 4.971e-10 140-155 PR00320C 13.01 9.280e-10 140-155
194	BL00678	Trp-Asp (WD) repeat proteins proteins.	BL00678 9.67 7.632e-09 142-153
196	PR00832	PAXILLIN SIGNATURE	PR00832B 9.87 9.174e-10 309-333
196	BL01160	Kinesin light chain repeat proteins.	BL01160B 19.54 2.054e-10 376-430 BL01160B 19.54 6.919e-10 383-437 BL01160B 19.54 9.676e-10 369-423
196	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 8.780e-09 40-55
196	BL00087	Copper/Zinc superoxide dismutase proteins.	BL00087C 20.18 8.784e-09 260-296
196	PR00806	VINCULIN SIGNATURE	PR00806A 6.63 9.014e-09 308-319
196	BL00326	Tropomyosins proteins.	BL00326A 14.01 9.143e-09 506-540
197	PR00674	LIGHT HARVESTING PROTEIN B CHAIN SIGNATURE	PR00674A 20.10 7.391e-09 134-155
198	PR00192	F-ACTIN CAPPING PROTEIN BETA SUBUNIT SIGNATURE	PR00192C 6.65 2.500e-36 57-84 PR00192D 8.23 4.462e-36 97-125 PR00192E 8.85 7.000e-33 212-239 PR00192A 8.23 1.474e-27 5-26 PR00192B 6.20 3.000e-26 26-48
198	BL00231	F-actin capping protein beta subunit proteins.	BL00231A 8.59 1.000e-40 5-51 BL00231B 14.16 1.000e-40 84-128 BL00231D 15.40 1.000e-40 165-200 BL00231E 11.66 1.000e-40 209-246 BL00231C 12.77 1.180e-15 146-157
199	PF00023	Ank repeat proteins.	PF00023A 16.03 4.750e-10 45-61
199	PF00791	Domain present in ZO-1 and Unc5-like netrin receptors.	PF00791B 28.49 8.768e-12 87-142 PF00791B 28.49 7.028e-09 499-116
199	BL01160	Kinesin light chain repeat proteins.	BL01160E 8.74 7.398e-09 323-362
201	PR00239	MOLLUSCAN RHODOPSIN C-TERMINAL TAIL SIGNATURE	PR00239E 1.58 6.114e-09 183-195
202	BL00412	Neuromodulin (GAP-43) proteins.	BL00412D 16.54 4.033e-10 319-370
202	BL00224	Clathrin light chain proteins.	BL00224B 16.94 4.845e-09 313-366
202	PF00992	Troponin.	PF00992A 16.67 8.734e-12 333-368 PF00992A 16.67 2.776e-09 344-379 PF00992A 16.67 5.026e-09 351-386
203	BL00790	Receptor tyrosine kinase class V proteins.	BL00790R 16.20 7.677e-09 29-73
204	BL00790	Receptor tyrosine kinase class V proteins.	BL00790R 16.20 7.677e-09 29-73
205	BL00790	Receptor tyrosine kinase class V proteins.	BL00790R 16.20 7.677e-09 29-73
207	BL00211	ABC transporters family proteins.	BL00211B 13.37 3.077e-17 573-167 BL00211B 13.37 7.577e-17 1204-1674

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
			BL00211A 12.23 1.900e-09 472-484
207	PR00478	PHOSPHORIBULOKINASE FAMILY SIGNATURE	PR00478A 13.44 4.133e-09 474-492
207	PR00802	SERUM ALBUMIN FAMILY SIGNATURE	PR00802G 14.57 7.188e-09 971-994
207	PR00836	SOMATOTROPIN HORMONE FAMILY SIGNATURE	PR00836D 13.05 7.125e-09 1504-1519
209	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 1.786e-10 288-303
210	BL00972	Ubiquitin carboxyl-terminal hydrolases family 2 proteins.	BL00972D 22.55 3.348e-11 388-413 BL00972E 20.72 4.343e-09 415-437
210	PR00198	ANNEXIN TYPE II SIGNATURE	PR00198H 12.05 7.750e-09 682-696
214	PD00469	PROTEIN PRECURSOR SIGNAL HYDROLA.	PD00469A 13.95 6.400e-09 73-86
215	PF00023	Ank repeat proteins.	PF00023A 16.03 8.875e-10 839-855 PF00023A 16.03 2.286e-09 884-900
215	PR00342	RHESUS BLOOD GROUP PROTEIN SIGNATURE	PR00342H 7.61 9.703e-09 317-340
217	BL00982	Bacterial-type phytoene dehydrogenase proteins.	BL00982A 18.41 8.013e-12 328-360
217	PR00368	FAD-DEPENDENT PYRIDINE NUCLEOTIDE REDUCTASE SIGNATURE	PR00368C 15.74 8.962e-11 326-352
217	PR00469	PYRIDINE NUCLEOTIDE DISULPHIDE REDUCTASE CLASS-II SIGNATURE	PR00469I 13.83 7.532e-11 449-468 PR00469F 16.51 7.152e-09 322-347
217	PD02042	IRON-SULFUR ELECTRON TRANSPORT AROMATIC HYDROCARB.	PD02042B 16.75 5.673e-09 126-141 PD02042A 21.13 9.045e-09 93-120
217	PR00419	ADRENODOXIN REDUCTASE FAMILY SIGNATURE	PR00419A 14.89 9.486e-09 326-349 PR00419D 10.62 9.534e-09 327-342
218	PF00157	PDZ domain proteins (Also known as DHR or GLGF).	PF00157 13.40 4.600e-09 688-699
219	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 7.000e-23 65-96 BL00107B 13.31 4.214e-10 130-146
219	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109B 12.27 7.102e-10 65-84
219	BL00240	Receptor tyrosine kinase class III proteins.	BL00240E 11.56 5.029e-09 51-89
220	PR00239	MOLLUSCAN RHODOPSIN C-TERMINAL TAIL SIGNATURE	PR00239E 1.58 3.045e-09 38-50
220	DM01803	1 HERPESVIRUS GLYCOPROTEIN H.	DM01803A 10.51 9.349e-09 34-55
220	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 5.160e-11 40-55 PR00049D 0.00 7.807e-11 41-56 PR00049D 0.00 8.336e-11 38-53 PR00049D 0.00 2.286e-10 42-57 PR00049D 0.00 8.857e-10 33-48 PR00049D 0.00 2.983e-09 37-52 PR00049D 0.00 9.847e-09 43-58
222	BL00326	Tropomyosins proteins.	BL00326A 14.01 5.337e-10 825-859

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
222	BL01160	Kinesin light chain repeat proteins.	BL01160B 19.54 9.924e-09 516-132
224	BL00478	LIM domain proteins.	BL00478B 14.79 8.527e-09 143-158
226	BL00048	Protamine P1 proteins.	BL00048 6.39 6.063e-09 199-226
228	BL00115	Eukaryotic RNA polymerase II heptapeptide repeat proteins.	BL00115Z 3.12 5.744e-10 113-162 BL00115Z 3.12 3.449e-09 120-169
228	BL00415	Synapsins proteins.	BL00415Q 2.23 8.723e-09 253-289
229	BL01161	Glucosamine/galactosamine-6-phosphate isomerases proteins.	BL01161A 19.47 1.000e-40 37-77 BL01161D 28.14 1.000e-40 199-244 BL01161B 21.37 5.091e-39 117-160 BL01161C 18.47 1.500e-23 170-199
231	PR00269	PLEIOTROPHIN/MIDKINE FAMILY SIGNATURE	PR00269A 13.91 3.133e-30 88-113
231	BL00181	PTN/MK heparin-binding protein family proteins.	BL00181A 19.07 4.960e-37 76-112 BL00181A 19.07 9.224e-18 78-114
236	BL00888	Cyclic nucleotide-binding domain proteins.	BL00888B 14.79 9.069e-13 499-523
236	BL00415	Synapsins proteins.	BL00415N 4.29 2.774e-09 733-777
236	PD00306	PROTEIN GLYCOPROTEIN PRECURSOR RE.	PD00306A 10.26 3.133e-09 646-660
236	PR00209	ALPHA/BETA GLIADIN FAMILY SIGNATURE	PR00209B 4.88 3.813e-09 739-758
236	DM00668	ZEIN.	DM00668A 10.20 8.500e-09 258-273
238	BL01188	GNS1/SUR4 family proteins.	BL01188B 13.46 4.115e-26 120-151 BL01188C 22.65 4.136e-26 151-202 BL01188D 8.62 1.290e-11 238-255 BL01188A 18.82 6.718e-10 55-87
239	PR00929	AT-HOOK-LIKE DOMAIN SIGNATURE	PR00929B 4.38 8.875e-09 133-583 PR00929C 5.26 8.914e-09 133-144
242	BL00232	Cadherins extracellular repeat proteins domain proteins.	BL00232B 32.79 2.765e-25 541-151 BL00232B 32.79 8.263e-22 766-814 BL00232B 32.79 2.397e-21 67-115 BL00232B 32.79 4.133e-19 1481-1529 BL00232B 32.79 1.000e-18 1371-1419 BL00232B 32.79 2.662e-18 1691-1739 BL00232B 32.79 5.292e-18 1287-1335 BL00232B 32.79 9.147e-18 1148-1196 BL00232B 32.79 1.265e-17 980-1028 BL00232B 32.79 1.529e-17 426-474 BL00232B 32.79 2.588e-17 1084-1132 BL00232B 32.79 1.386e-16 1184-1232 BL00232C 10.65 5.390e-12 1369-1387 BL00232C 10.65 1.391e-11 204-660 BL00232C 10.65 2.174e-11 1584-1164 BL00232C 10.65 4.522e-11 1689-1707 BL00232C 10.65 1.000e-10 65-83 BL00232C 10.65 4.115e-10 1285-1303 BL00232B 32.79 7.200e-10 649-697 BL00232C 10.65 9.827e-10 978-996 BL00232C 10.65 1.947e-09 170-188 BL00232B 32.79 2.137e-09 172-220

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
			BL00232C 10.65 4.474e-09 1182-1200 BL00232C 10.65 8.737e-09 539-119
243	BL00795	Involucrin proteins.	BL00795C 17.06 4.977e-10 64-109 BL00795C 17.06 6.300e-09 55-100
244	BL00790	Receptor tyrosine kinase class V proteins.	BL00790I 20.01 7.823e-15 23-54 BL00790I 20.01 9.400e-11 310-341 BL00790I 20.01 1.900e-10 117-148 BL00790I 20.01 3.893e-09 215-246
244	PR00014	FIBRONECTIN TYPE III REPEAT SIGNATURE	PR00014D 12.04 6.400e-11 30-45 PR00014D 12.04 6.400e-11 317-332 PR00014C 15.44 9.171e-09 204-223 PR00014D 12.04 1.000e-08 222-237
245	BL00183	Ubiquitin-conjugating enzymes proteins.	BL00183 28.97 7.037e-10 140-188
246	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019A 11.19 8.800e-12 205-219 PR00019B 11.36 2.000e-11 202-216
247	BL00214	Cytosolic fatty-acid binding proteins.	BL00214B 26.51 7.180e-24 206-251 BL00214A 21.17 6.250e-22 165-191
247	PR00178	FATTY ACID-BINDING PROTEIN SIGNATURE	PR00178A 15.07 4.913e-21 166-187 PR00178C 20.54 2.500e-17 226-254 PR00178D 13.52 6.897e-16 272-291 PR00178B 10.52 4.900e-10 200-212
248	PR00395	RIBOSOMAL PROTEIN S2 SIGNATURE	PR00395C 16.17 2.047e-13 46-64
248	BL00962	Ribosomal protein S2 proteins.	BL00962C 15.90 2.846e-12 46-64
249	BL00227	Tubulin subunits alpha, beta, and gamma proteins.	BL00227D 18.46 1.000e-40 74-128 BL00227F 21.16 1.529e-33 226-280 BL00227E 24.15 1.409e-26 178-213
250	BL00227	Tubulin subunits alpha, beta, and gamma proteins.	BL00227C 25.48 1.000e-40 39-91 BL00227D 18.46 1.000e-40 148-202 BL00227F 21.16 1.529e-33 300-354 BL00227E 24.15 1.409e-26 252-287
251	BL00152	ATP synthase alpha and beta subunits proteins.	BL00152B 21.40 1.900e-31 191-229 BL00152A 15.38 5.154e-21 134-160 BL00152C 11.41 6.250e-12 291-303
252	BL00152	ATP synthase alpha and beta subunits proteins.	BL00152E 22.68 1.000e-32 285-323 BL00152A 15.38 5.154e-21 134-160 BL00152C 11.41 6.250e-12 247-259
253	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 2.200e-11 54-63
253	BL01282	BIR repeat proteins.	BL01282B 30.49 2.029e-09 35-74
254	DM00892	3 RETROVIRAL PROTEINASE.	DM00892C 23.55 9.739e-12 417-451
254	PR00417	PROKARYOTIC DNA TOPOISOMERASE I SIGNATURE	PR00417A 12.66 8.472e-09 65-79
255	BL01052	Calponin family repeat proteins.	BL01052C 18.51 1.000e-40 88-128 BL01052A 16.12 2.875e-35 3-35 BL01052B 15.31 5.219e-26 52-78
255	PR00888	SMOOTH MUSCLE PROTEIN/CALPONIN FAMILY SIGNATURE	PR00888D 16.09 9.112e-19 89-106 PR00888E 11.81 2.800e-18 105-121 PR00888F 7.44 4.600e-18 126-141

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
			PR00888A 11.87 7.750e-18 3-18 PR00888C 12.27 2.286e-17 52-68 PR00888G 12.73 9.438e-15 163-177 PR00888B 13.72 1.321e-14 22-36
255	PR00890	SMOOTH MUSCLE PROTEIN 22-ALPHA (TRANSGELIN) SIGNATURE	PR00890E 14.34 1.429e-27 136-156 PR00890A 8.61 1.000e-26 34-54 PR00890C 8.22 1.600e-19 85-99 PR00890B 8.75 6.318e-19 62-78 PR00890F 12.92 1.205e-17 162-175 PR00890D 16.17 1.130e-13 119-129
257	BL00745	Prokaryotic-type class I peptide chain release factors signat.	BL00745C 13.66 1.000e-40 202-249 BL00745B 22.56 8.683e-33 148-191 BL00745D 14.90 8.435e-23 280-303
259	BL00194	Thioredoxin family proteins.	BL00194 12.16 7.429e-10 684-697
260	BL00612	Osteonectin domain proteins.	BL00612E 13.12 3.948e-10 391-436
260	BL00484	Thyroglobulin type-1 repeat proteins proteins.	BL00484C 17.01 8.244e-11 136-151 BL00484B 9.04 2.145e-10 249-263 BL00484C 17.01 2.309e-09 269-284 BL00484B 9.04 8.950e-09 116-130
262	PR00187	DNAJ PROTEIN FAMILY SIGNATURE	PR00187A 12.84 2.375e-09 288-308
262	BL00198	Nt-dnaJ domain proteins.	BL00198A 8.07 3.681e-09 292-309
262	BL00157	Aminotransferases class-V pyridoxal-phosphate attachment site proteins.	BL00157A 11.72 8.200e-09 16-26
263	PR00320	G-PROTEIN BETA WD-40 REPEAT SIGNATURE	PR00320B 12.19 2.125e-09 207-222
263	PF00913	Trypanosome variant surface glycoprotein.	PF00913A 7.33 2.500e-09 666-673
266	BL01144	Ribosomal protein L31e proteins.	BL01144 25.07 1.000e-40 21-73
268	DM00516	186 DISCOIDIN I N-TERMINAL.	DM00516 30.53 8.168e-13 153-198
268	BL00132	Zinc carboxypeptidases, zinc-binding region 1 proteins.	BL00132C 21.35 7.863e-10 307-348 BL00132A 26.07 8.988e-10 224-265
268	PR00765	CARBOXYPEPTIDASE A METALLOPROTEASE (M14) FAMILY SIGNATURE	PR00765B 15.57 7.171e-12 276-291 PR00765D 14.16 1.551e-09 420-434
268	BL00170	Cyclophilin-type peptidyl-prolyl cis-trans isomerase signatur.	BL00170A 17.08 9.018e-09 485-512
269	BL00622	Bacterial regulatory proteins, luxR family proteins.	BL00622 32.69 9.780e-09 11-58
270	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 1.000e-11 447-461 PR00048A 10.52 4.316e-11 389-403 PR00048A 10.52 6.684e-11 362-376
270	PF00651	BTB (also known as BR-C/Ttk) domain proteins.	PF00651 15.00 3.143e-10 37-50
270	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 7.000e-10 392-409 BL00028 16.07 9.100e-10 256-273 BL00028 16.07 2.286e-09 450-467 BL00028 16.07 8.714e-09 365-382
274	DM00303	6 LEA 11-MER REPEAT REPEAT.	DM00303A 13.20 3.310e-09 467-517
275	PF00622	Domain in SPLa and the RYanodine	PF00622B 21.00 9.357e-14 374-396

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
		Receptor.	PF00622C 12.62 1.857e-12 458-472
275	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 8.800e-11 44-53
277	PF00651	BTB (also known as BR-C/Ttk) domain proteins.	PF00651 15.00 9.133e-10 65-78
278	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 8.200e-16 295-308 PD00066 13.92 8.200e-16 519-532 PD00066 13.92 1.692e-15 351-364 PD00066 13.92 4.462e-15 547-122 PD00066 13.92 4.600e-14 323-336 PD00066 13.92 4.600e-14 435-448 PD00066 13.92 7.000e-14 463-476 PD00066 13.92 1.500e-13 239-252 PD00066 13.92 3.143e-12 267-280 PD00066 13.92 3.143e-12 407-420 PD00066 13.92 8.826e-11 211-224 PD00066 13.92 2.038e-10 491-504 PD00066 13.92 2.385e-10 379-392
278	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 7.750e-16 444-458 PR00048A 10.52 6.727e-15 360-374 PR00048A 10.52 9.182e-15 528-542 PR00048A 10.52 7.000e-14 472-486 PR00048A 10.52 7.750e-14 388-402 PR00048A 10.52 1.000e-13 332-346 PR00048A 10.52 3.133e-13 304-318 PR00048A 10.52 4.857e-13 118-132 PR00048A 10.52 6.786e-13 500-514 PR00048B 6.02 1.000e-12 292-302 PR00048A 10.52 8.941e-12 192-206 PR00048B 6.02 1.000e-11 348-358 PR00048A 10.52 1.947e-11 248-262 PR00048B 6.02 2.385e-11 264-274 PR00048B 6.02 7.231e-11 544-116 PR00048A 10.52 7.632e-11 416-430 PR00048B 6.02 8.615e-11 236-246 PR00048B 6.02 2.688e-10 516-526 PR00048B 6.02 4.375e-10 460-470 PR00048B 6.02 4.375e-10 488-498 PR00048B 6.02 4.938e-10 404-414 PR00048B 6.02 6.063e-10 320-330 PR00048A 10.52 7.214e-10 220-234 PR00048B 6.02 1.947e-09 432-442 PR00048B 6.02 4.316e-09 572-144
278	DM01970	0 kw ZK632.12 YDR313C ENDOSOMAL III.	DM01970B 8.60 5.012e-09 191-204
279	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 6.400e-16 449-462 PD00066 13.92 6.538e-15 504-517 PD00066 13.92 9.308e-15 421-434 PD00066 13.92 7.000e-14 476-489 PD00066 13.92 6.087e-11 393-406
279	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 2.500e-17 350-367 BL00028 16.07 5.050e-13 405-422 BL00028 16.07

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
			9.171e-12 433-450 BL00028 16.07 2.731e-11
			488-505 BL00028 16.07 3.077e-11 516-533 BL00028 16.07 6.100e-10 377-394
279	PD02462	PROTEIN BOLA TRANSCRIPTION REGULATION AC.	PD02462A 22.48 6.488e-09 481-516
279	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 3.250e-16 347-361 PR00048B 6.02 5.154e-11 501-511 PR00048B 6.02 1.000e-10 446-456 PR00048A 10.52 1.391e-10 513-527 PR00048A 10.52 2.565e-10 485-499 PR00048A 10.52 5.696e-10 402-416 PR00048B 6.02 8.875e-10 418-428 PR00048A 10.52 1.720e-09 430-444 PR00048B 6.02 3.368e-09 390-400 PR00048A 10.52 8.200e-09 374-388
285	BL00276	Channel forming colicins proteins.	BL00276A 8.87 6.500e-09 257-269
286	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 2.000e-30 10-49
286	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 6.400e-16 388-401 PD00066 13.92 3.769e-15 248-261 PD00066 13.92 9.308e-15 304-317 PD00066 13.92 2.200e-14 360-373 PD00066 13.92 2.200e-14 416-429 PD00066 13.92 6.400e-14 332-345 PD00066 13.92 1.000e-13 220-233 PD00066 13.92 2.500e-13 192-205 PD00066 13.92 5.000e-13 276-289 PD00066 13.92 5.500e-09 136-149
286	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 2.286e-16 260-277 BL00028 16.07 2.588e-14 288-305 BL00028 16.07 2.800e-13 400-417 BL00028 16.07 6.850e-13 120-137 BL00028 16.07 3.423e-11 148-165 BL00028 16.07 7.923e-11 344-361 BL00028 16.07 2.500e-10 204-221 BL00028 16.07 2.500e-10 428-445 BL00028 16.07 3.100e-10 316-333 BL00028 16.07 6.100e-10 176-193 BL00028 16.07 1.771e-09 232-249 BL00028 16.07 8.200e-09 372-389
286	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 7.000e-17 257-271 PR00048A 10.52 6.727e-15 397-411 PR00048A 10.52 2.929e-13 285-299 PR00048A 10.52 9.471e-12 369-383 PR00048B 6.02 1.000e-11 329-339 PR00048A 10.52 1.474e-11 313-327 PR00048A 10.52 2.421e-11 425-439 PR00048B 6.02 3.077e-11 385-395 PR00048A 10.52 6.684e-11 117-131 PR00048A 10.52 8.141e-11 201-215 PR00048A 10.52 1.783e-10 341-355 PR00048B 6.02 2.125e-10 301-311

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
			PR00048B 6.02 2.125e-10 357-367 PR00048B 6.02 2.688e-10 217-227 PR00048A 10.52 3.739e-10 229-243 PR00048B 6.02 4.938e-10 273-283 PR00048B 6.02 1.474e-09 245-255 PR00048A 10.52 2.440e-09 145-159 PR00048B 6.02 3.842e-09 161-171 PR00048B 6.02 8.105e-09 441-451 PR00048B 6.02 9.053e-09 189-199
287	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 7.407e-23 3-42
287	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 8.941e-14 269-286 BL00028 16.07 1.000e-13 549-128 BL00028 16.07 2.565e-12 194-650 BL00028 16.07 6.087e- 12 241-258 BL00028 16.07 6.870e-12 297- 314 BL00028 16.07 6.870e-12 381-398 BL00028 16.07 7.214e-12 493-510 BL00028 16.07 1.346e-11 465-482 BL00028 16.07 1.692e-11 353-370 BL00028 16.07 3.769e- 11 325-342 BL00028 16.07 6.192e-11 167- 622 BL00028 16.07 8.962e-11 213-230 BL00028 16.07 1.600e-10 409-426 BL00028 16.07 5.200e-10 185-202 BL00028 16.07 6.700e-10 577-156 BL00028 16.07 3.057e- 09 521-538 BL00028 16.07 6.143e-09 437- 454
287	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 9.250e-14 238-252 PR00048A 10.52 3.209e-12 266-280 PR00048A 10.52 4.706e-12 490-504 PR00048A 10.52 5.765e-12 462-476 PR00048A 10.52 7.882e-12 630-644 PR00048A 10.52 8.941e-12 518-532 PR00048A 10.52 9.471e-12 164-178 PR00048A 10.52 5.737e-11 378-392 PR00048A 10.52 7.158e-11 546-122 PR00048B 6.02 7.231e-11 180-190 PR00048A 10.52 8.141e-11 210-224 PR00048A 10.52 9.053e-11 294-308 PR00048A 10.52 9.053e-11 406-420 PR00048A 10.52 3.348e-10 322-336 PR00048B 6.02 3.813e-10 338-348 PR00048B 6.02 3.813e-10 394-404 PR00048B 6.02 3.813e-10 478-488 PR00048B 6.02 4.938e-10 506-516 PR00048A 10.52 8.043e-10 434-448 PR00048B 6.02 8.875e-10 226-236 PR00048B 6.02 8.875e-10 450-460 PR00048B 6.02 1.000e-09 366-376 PR00048B 6.02 1.000e-09 422-432 PR00048A 10.52 3.520e-09 136-588

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
			PR00048B 6.02 7.158e-09 590-600 PR00048B 6.02 7.632e-09 310-320 PR00048B 6.02 7.632e-09 124-572 PR00048A 10.52 9.280e-09 350-364
289	PR00070	DIHYDROFOLATE REDUCTASE SIGNATURE	PR00070C 13.09 6.143e-16 51-63 PR00070D 11.63 2.929e-15 112-127
289	BL00075	Dihydrofolate reductase proteins.	BL00075A 27.70 7.900e-16 8-39 BL00075B 13.49 3.813e-15 51-63 BL00075C 8.51 2.862e-11 66-79 BL00075D 5.74 8.105e-10 113-123
292	PR00250	FUNGAL PHEROMONE MATING FACTOR-STE2 GPCR SIGNATURE	PR00250D 14.62 9.163e-09 254-278
294	PR00081	GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE	PR00081A 10.53 2.731e-09 39-57
294	PR00080	ALCOHOL DEHYDROGENASE SUPERFAMILY SIGNATURE	PR00080C 17.16 6.464e-11 191-211 PR00080A 9.32 9.750e-09 118-130
295	PR00806	VINCULIN SIGNATURE	PR00806B 4.28 8.920e-09 276-290 PR00806B 4.28 9.202e-09 275-289
296	PF00992	Troponin.	PF00992A 16.67 3.789e-10 553-588
296	BL00752	XPA protein.	BL00752B 19.17 8.144e-09 130-612
296	BL01160	Kinesin light chain repeat proteins.	BL01160B 19.54 8.551e-09 536-590
298	PR00511	TEKTIN SIGNATURE	PR00511C 7.86 4.214e-09 371-388
300	BL00353	HMG1/2 proteins.	BL00353B 11.47 9.171e-19 228-278
301	PR00240	ALPHA-1A ADRENERGIC RECEPTOR SIGNATURE	PR00240C 8.38 3.941e-10 316-336
302	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 2.200e-11 54-63
302	BL01282	BIR repeat proteins.	BL01282B 30.49 2.029e-09 35-74
305	PR00193	MYOSIN HEAVY CHAIN SIGNATURE	PR00193D 14.36 1.545e-31 390-419 PR00193C 12.60 1.209e-25 143-171 PR00193B 11.69 2.543e-24 95-121 PR00193A 15.41 6.885e-19 39-59 PR00193E 19.47 3.291e-12 444-473
305	BL00675	Sigma-54 interaction domain proteins ATP-binding region A proteins.	BL00675A 24.86 3.475e-09 98-142
306	PR00239	MOLLUSCAN RHODOPSIN C-TERMINAL TAIL SIGNATURE	PR00239E 1.58 5.920e-11 47-59
306	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 7.923e-15 140-153 PD00066 13.92 4.000e-14 112-125 PD00066 13.92 1.391e-11 84-97 PD00066 13.92 1.692e-10 168-181
306	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 2.059e-14 96-113 BL00028 16.07 4.130e-12 124-141 BL00028 16.07 2.385e-11 68-85 BL00028 16.07 8.269e-11 180-197 BL00028 16.07 8.962e-11 152-169 BL00028 16.07 9.400e-10 319-336
306	PR00799	ASPARTATE AMINOTRANSFERASE SIGNATURE	PR00799D 16.46 5.125e-09 188-214

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
306	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048B 6.02 1.900e-13 81-91 PR00048A 10.52 3.133e-13 65-79 PR00048A 10.52 9.357e-13 121-135 PR00048A 10.52 9.357e-13 149-163 PR00048B 6.02 2.688e-10 137-147 PR00048A 10.52 4.522e-10 279-293 PR00048A 10.52 5.696e-10 177-191 PR00048B 6.02 9.438e-10 109-119 PR00048A 10.52 3.160e-09 93-107 PR00048B 6.02 8.105e-09 165-175
307	PD00015	GLYCOPROTEIN PRECURSOR CELL SI.	PD00015A 8.90 6.400e-09 35-43
310	DM00031	IMMUNOGLOBULIN V REGION.	DM00031B 15.41 3.662e-11 80-114
311	BL00824	Elongation factor 1 beta/beta/delta chain proteins.	BL00824C 14.58 1.000e-40 129-167 BL00824D 14.04 6.192e-39 167-202 BL00824B 9.21 2.080e-21 96-116 BL00824E 12.49 3.333e-19 210-226
312	PR00501	KELCH REPEAT SIGNATURE	PR00501B 18.88 7.632e-09 476-491 PR00501B 18.88 9.763e-09 523-538
313	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 6.200e-30 43-82
313	PD00066	PROTEIN ZINC-FINGER METAL-. BINDI	PD00066 13.92 6.500e-13 439-452 PD00066 13.92 8.000e-13 355-368 PD00066 13.92 1.000e-12 383-396 PD00066 13.92 4.000e-12 327-340 PD00066 13.92 5.714e-12 411-424 PD00066 13.92 8.435e-11 299-312 13.92 5.800e-14 467-480 PD00066
313	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 2.565e-12 451-468 BL00028 16.07 2.957e-12 311-328 BL00028 16.07 3.348e-12 367-384 BL00028 16.07 1.692e-11 423-440 BL00028 16.07 2.731e-11 283-300 BL00028 16.07 2.800e-10 339-356 BL00028 16.07 9.700e-10 199-216 BL00028 16.07 1.000e-09 395-412 BL00028 16.07 4.086e-09 120-137
313	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 5.909e-15 364-378 PR00048A 10.52 2.286e-13 308-322 PR00048A 10.52 7.429e-13 392-406 PR00048A 10.52 6.824e-12 448-462 PR00048A 10.52 2.421e-11 196-210 PR00048A 10.52 1.000e-10 280-294 PR00048B 6.02 3.813e-10 324-334 PR00048B 6.02 4.375e-10 464-474 PR00048A 10.52 6.870e-10 336-350 PR00048A 10.52 7.214e-10 420-434 PR00048B 6.02 7.750e-10 436-446 PR00048B 6.02 4.316e-09 380-390
314	PR00121	SODIUM/POTASSIUM-TRANSPORTING ATPASE SIGNATURE	PR00121D 16.72 1.577e-13 210-232
314	PR00119	P-TYPE CATION-TRANSPORTING	PR00119B 13.94 9.194e-12 217-232

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
		ATPASE SUPERFAMILY SIGNATURE	
314	BL01228	Hypothetical cof family proteins.	BL01228D 17.44 3.400e-11 646-671
314	BL00154	E1-E2 ATPases phosphorylation site proteins.	BL00154E 20.37 4.054e-13 486-527 BL00154C 12.38 4.060e-12 213-232 BL00154F 8.23 9.597e-11 207-669
315	BL00888	Cyclic nucleotide-binding domain proteins.	BL00888B 14.79 1.692e-10 396-420
315	BL00420	Speract receptor repeat proteins domain proteins.	BL00420A 20.42 8.338e-09 215-682
315	DM00668	ZEIN.	DM00668A 10.20 8.500e-09 155-170
316	PR00727	BACTERIAL LEADER PEPTIDASE I (S26) FAMILY SIGNATURE	PR00727C 13.04 9.063e-16 108-128 PR00727B 12.51 7.848e-11 81-94
316	BL00501	Signal peptidases I serine proteins.	BL00501D 16.69 2.884e-13 108-128 BL00501C 9.61 9.561e-11 81-93 BL00501B 12.58 7.000e-09 61-77
317	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 9.471e-27 13-52
317	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 5.235e-14 214-231 BL00028 16.07 6.850e-13 270-287 BL00028 16.07 9.100e-13 354-371 BL00028 16.07 1.391e-12 158-175 BL00028 16.07 1.346e-11 298-315 BL00028 16.07 3.769e-11 242-259 BL00028 16.07 6.538e-11 380-397 BL00028 16.07 8.800e-10 186-203 BL00028 16.07 1.514e-09 326-343
317	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048B 6.02 3.000e-12 199-209 PR00048A 10.52 7.882e-12 351-365 PR00048A 10.52 8.412e-12 323-337 PR00048A 10.52 8.941e-12 239-253 PR00048A 10.52 1.474e-11 211-225 PR00048A 10.52 6.211e-11 155-169 PR00048B 6.02 7.231e-11 311-321 PR00048A 10.52 8.141e-11 267-281 PR00048B 6.02 3.250e-10 339-349 PR00048B 6.02 3.813e-10 255-265 PR00048B 6.02 7.188e-10 283-293 PR00048B 6.02 3.842e-09 171-181 PR00048B 6.02 3.842e-09 393-403 PR00048A 10.52 8.200e-09 295-309
319	PR00004	ANAPHYLATOXIN DOMAIN SIGNATURE	PR00004C 12.46 8.141e-09 91-103
320	DM00060	338 kw NEUREXIN ALPHA III CYSTEINE.	DM00060 6.92 6.500e-11 28-38
320	PR00010	TYPE II EGF-LIKE SIGNATURE	PR00010C 11.16 7.667e-11 44-55
325	PR00020	MAM DOMAIN SIGNATURE	PR00020A 18.17 5.776e-12 344-363 PR00020C 13.66 6.932e-10 417-429
325	BL00740	MAM domain proteins.	BL00740A 13.87 8.313e-12 346-359 BL00740B 19.76 8.500e-09 486-507
325	PD02080	T-CELL GLYCOPROTEIN CD8	PD02080B 20.69 9.621e-09 123-162

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
		CHAIN SURFACE ALPHA PRE.	
326	BL00048	Protamine P1 proteins.	BL00048 6.39 6.128e-10 167-194
326	PF01140	Matrix protein (MA), p15.	PF01140D 15.54 9.791e-09 220-255
327	PR00020	MAM DOMAIN SIGNATURE	PR00020C 13.66 2.615e-11 143-593 PR00020B 15.52 5.059e-10 52-69 PR00020B 15.52 1.789e-09 553-132
329	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 9.357e-32 8-47
329	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 3.209e-14 284-301 BL00028 16.07 4.600e-13 508-525 BL00028 16.07 6.400e-13 368-385 BL00028 16.07 4.115e-11 396-413 BL00028 16.07 4.115e-11 424-441 BL00028 16.07 8.269e-11 172-189 BL00028 16.07 8.962e-11 256-273 BL00028 16.07 9.308e-11 312-329 BL00028 16.07 9.654e-11 200-217 BL00028 16.07 3.100e-10 340-357 BL00028 16.07 5.500e-10 452-469 BL00028 16.07 9.100e-10 480-497 BL00028 16.07 4.086e-09 228-245
329	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 7.000e-14 272-285 PD00066 13.92 5.000e-13 328-341 PD00066 13.92 5.500e-13 188-201 PD00066 13.92 5.500e-13 384-397 PD00066 13.92 6.000e-13 496-509 PD00066 13.92 6.143e-12 468-481 PD00066 13.92 2.731e-10 440-453 PD00066 13.92 4.808e-10 160-173 PD00066 13.92 5.500e-10 244-257 PD00066 13.92 7.000e-09 216-229 PD00066 13.92 7.000e-09 412-425
332	PD02870	RECEPTOR INTERLEUKIN-1 PRECURSOR.	PD02870B 18.83 5.871e-11 468-501
332	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019A 11.19 8.043e-10 275-289
332	BL00240	Receptor tyrosine kinase class III proteins.	BL00240B 24.70 4.447e-09 430-454
333	BL00738	S-adenosyl-L-homocysteine hydrolase proteins.	BL00738J 18.61 1.000e-40 154-204 BL00738H 23.08 5.320e-36 468-521 BL00738F 12.23 7.261e-29 387-419 BL00738A 16.27 9.660e-27 216-256 BL00738C 16.53 7.923e-25 281-319 BL00738G 14.29 6.268e-23 446-468 BL00738B 12.28 8.085e-21 256-281 BL00738E 14.18 9.200e-19 361-384 BL00738I 14.57 5.135e-17 545-583 BL00738D 7.16 5.109e-13 335-350
333	BL00836	Alanine dehydrogenase & pyridine nucleotide transhydrogenase.	BL00836D 22.30 8.622e-09 424-461
337	PR00425	BRADYKININ RECEPTOR SIGNATURE	PR00425C 13.23 3.148e-09 80-100
342	PD01823	PROTEIN INTERGENIC REGION	PD01823E 9.30 6.824e-12 108-121

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
		ABC1 PRECURSOR MITOCHONDRION T.	PD01823D 16.66 1.265e-09 46-67
343	PR00976	RIBOSOMAL PROTEIN S21 FAMILY SIGNATURE.	PR00976C 10.41 2.837e-09 396-407
343	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 1.458e-09 473-506 DM00215 19.43 4.814e-09 463-496
343	PR00671	INHIBIN BETA B CHAIN SIGNATURE	PR00671C 4.18 9.172e-09 707-727
343	PD01234	PROTEIN NUCLEAR BROMODOMAIN TRANS.	PD01234B 15.53 1.000e-08 482-500
344	PR00175	MYOGLOBIN SIGNATURE	PR00175B 9.02 2.143e-10 25-49
344	PR00814	BETA HAEMOGLOBIN SIGNATURE	PR00814C 9.20 6.523e-10 66-84
344	PR00173	ERYTHROCRUORIN FAMILY SIGNATURE	PR00173A 15.91 7.158e-10 25-48
344	BL01033	Globins profile.	BL01033A 16.94 1.000e-16 25-47 BL01033B 13.81 8.615e-09 87-99
344	PR00612	ALPHA HAEMOGLOBIN SIGNATURE	PR00612E 9.04 4.194e-12 122-139 PR00612B 10.92 3.483e-10 32-43 PR00612D 9.76 9.438e-09 74-88
345	PR00814	BETA HAEMOGLOBIN SIGNATURE	PR00814C 9.20 6.523e-10 104-122
345	BL01033	Globins profile.	BL01033A 16.94 5.125e-10 63-85 BL01033B 13.81 8.615e-09 125-137
345	PR00612	ALPHA HAEMOGLOBIN SIGNATURE	PR00612E 9.04 4.194e-12 160-177 PR00612B 10.92 3.483e-10 70-81 PR00612D 9.76 9.438e-09 112-126
349	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 6.133e-32 6-45
350	BL00972	Ubiquitin carboxyl-terminal hydrolases family 2 proteins.	BL00972A 11.93 6.318e-19 364-382 BL00972D 22.55 7.968e-16 210-673 BL00972B 9.45 1.600e-12 445-455
350	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 8.008e-13 121-136 PR00049D 0.00 7.375e-12 125-140 PR00049D 0.00 5.916e-11 128-143 PR00049D 0.00 6.748e-11 122-137 PR00049D 0.00 9.395e-11 126-141 PR00049D 0.00 1.286e-10 119-134 PR00049D 0.00 8.929e-10 127-142 PR00049D 0.00 2.678e-09 129-144 PR00049D 0.00 4.051e-09 123-138 PR00049D 0.00 4.051e-09 124-139 PR00049D 0.00 4.051e-09 130-145
350	PR00211	GLUTELIN SIGNATURE	PR00211B 0.86 7.500e-09 124-145
350	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 5.339e-10 108-141 DM00215 19.43 7.268e-10 112-145 DM00215 19.43 2.525e-09 106-139 DM00215 19.43 9.695e-09 107-140
350	BL00048	Protamine P1 proteins.	BL00048 6.39 9.888e-09 145-172
352	BL00518	Zinc finger, C3HC4 type (RING	BL00518 12.23 4.429e-10 214-223

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
		finger), proteins.	
353	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 4.429e-10 179-188
354	BL01009	Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 proteins.	BL01009D 14.19 9.341e-17 160-181 BL01009A 13.75 3.769e-14 80-98 BL01009E 13.50 5.333e-14 194-210 BL01009C 10.54 2.667e-11 127-141
354	PR00838	VENOM ALLERGEN 5 SIGNATURE	PR00838G 16.07 2.304e-14 158-178 PR00838D 8.73 4.452e-12 80-99 PR00838F 10.11 7.532e-10 125-141
354	PR00837	ALLERGEN V5/TPX-1 FAMILY SIGNATURE	PR00837C 17.21 7.429e-18 159-176 PR00837A 14.77 1.900e-15 80-99 PR00837D 11.12 2.198e-13 195-209 PR00837B 11.64 3.483e-09 127-141
356	BL00215	Mitochondrial energy transfer proteins.	BL00215A 15.82 8.500e-17 16-41 BL00215B 10.44 4.900e-09 177-190 BL00215A 15.82 6.786e-09 133-158 BL00215B 10.44 7.300e-09 278-291
356	PR00926	MITOCHONDRIAL CARRIER PROTEIN SIGNATURE	PR00926E 11.70 6.049e-13 91-110 PR00926F 17.75 7.600e-11 240-263 PR00926F 17.75 5.219e-10 18-41 PR00926D 10.53 7.387e-09 246-265
357	PR00326	GTP1/OBG GTP-BINDING PROTEIN FAMILY SIGNATURE	PR00326A 8.75 7.150e-11 21-42
357	BL00113	Adenylate kinase proteins.	BL00113A 12.74 6.677e-09 22-39
357	BL01128	Shikimate kinase proteins.	BL01128A 18.84 7.802e-09 21-55
357	BL00300	SRP54-type proteins GTP-binding domain proteins.	BL00300B 20.56 1.000e-08 18-64
358	BL00972	Ubiquitin carboxyl-terminal hydrolases family 2 proteins.	BL00972A 11.93 6.318e-19 324-342 BL00972D 22.55 3.903e-16 170-194 BL00972B 9.45 1.600e-12 405-415
364	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 1.482e-10 355-388
364	PR00217	43 KD POSTSYNAPTIC PROTEIN SIGNATURE	PR00217C 10.91 4.600e-10 302-318
365	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 2.800e-11 125-134
365	BL00415	Synapsins proteins.	BL00415N 4.29 2.839e-09 387-431
365	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 7.706e-11 377-410 DM00215 19.43 8.412e-11 333-366 DM00215 19.43 2.678e-09 356-389 DM00215 19.43 5.138e-09 376-409
365	BL01102	Prokaryotic dksA/traR C4-type zinc finger.	BL01102 15.99 5.705e-09 109-135
365	PR00211	GLUTELIN SIGNATURE	PR00211B 0.86 5.959e-11 407-428 PR00211B 0.86 2.212e-10 401-422 PR00211B 0.86 9.500e-09 336-357
365	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 9.695e-09 335-350
367	PD00126	PROTEIN REPEAT DOMAIN TPR NUCLEA.	PD00126A 22.53 8.448e-09 2-23

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
370	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 7.353e-14 157-174 BL00028 16.07 1.000e-13 269-286 BL00028 16.07 8.200e-13 493-510 BL00028 16.07 3.739e-12 213-230 BL00028 16.07 6.478e-12 381-398 BL00028 16.07 1.346e-11 185-202 BL00028 16.07 2.385e-11 129-146 BL00028 16.07 2.385e-11 325-342 BL00028 16.07 5.154e-11 241-258 BL00028 16.07 9.654e-11 437-454 BL00028 16.07 1.300e-10 297-314 BL00028 16.07 9.100e-10 409-426 BL00028 16.07 9.100e-10 465-482
370	PD00066	PROTEIN ZINC-FINGER METAL-BINDL	PD00066 13.92 2.385e-15 229-242 PD00066 13.92 3.077e-15 145-158 PD00066 13.92 8.800e-14 173-186 PD00066 13.92 3.500e-13 369-382 PD00066 13.92 8.500e-13 341-354 PD00066 13.92 9.133e-12 397-410 PD00066 13.92 2.174e-11 313-326 PD00066 13.92 3.348e-11 453-466 PD00066 13.92 3.739e-11 481-494 PD00066 13.92 7.214e-11 257-270 PD00066 13.92 2.038e-10 425-438 PD00066 13.92 6.538e-10 201-214 PD00066 13.92 5.200e-09 285-298
370	DM01970	0 kw ZK632.12 YDR313C ENDOSOMAL III.	DM01970B 8.60 6.201e-09 265-278
370	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 1.474e-11 462-476 PR00048A 10.52 6.684e-11 182-196 PR00048A 10.52 2.957e-10 434-448 PR00048B 6.02 5.500e-10 338-348 PR00048A 10.52 6.478e-10 350-364 PR00048B 6.02 6.187e-10 226-236 PR00048A 10.52 6.870e-10 490-504 PR00048A 10.52 8.826e-10 406-420 PR00048B 6.02 3.842e-09 170-180 PR00048B 6.02 4.316e-09 366-376 PR00048B 6.02 4.789e-09 478-488 PR00048B 6.02 7.632e-09 142-152 PR00048A 10.52 8.122e-09 126-140 PR00048B 6.02 9.053e-09 450-460
371	BL01019	ADP-ribosylation factors family proteins.	BL01019B 19.49 6.276e-21 95-150 BL01019A 13.20 8.453e-17 51-91
371	PR00328	GTP-BINDING SAR1 PROTEIN SIGNATURE	PR00328C 13.16 8.481e-13 78-104 PR00328D 12.56 3.357e-11 123-145
371	BL01115	GTP-binding nuclear protein ran proteins.	BL01115A 10.22 8.119e-11 21-65
373	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 4.522e-12 208-225
373	PD00066	PROTEIN ZINC-FINGER METAL-BINDL	PD00066 13.92 7.000e-13 194-207 PD00066 13.92 7.000e-13 224-237 PD00066 13.92 7.000e-12 254-267
373	PR00048	C2H2-TYPE ZINC FINGER	PR00048A 10.52 1.391e-10 205-219

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
		SIGNATURE	PR00048B 6.02 6.063e-10 221-231
374	PR00308	TYPE I ANTIFREEZE PROTEIN SIGNATURE	PR00308A 5.90 7.288e-11 533-548 PR00308A 5.90 8.835e-09 534-549
377	PD02784	PROTEIN NUCLEAR RIBONUCLEOPROTEIN.	PD02784B 26.46 7.538e-09 147-190
378	PD01351	PROTEIN REPEAT NEUROFILAMENT TRIPL.	PD01351A 8.69 7.469e-09 155-166
380	PF00094	von Willebrand factor type D domain proteins.	PF00094C 12.88 1.918e-09 43-53
380	BL01208	VWFC domain proteins.	BL01208B 15.83 3.667e-11 120-135 BL01208B 15.83 1.973e-09 178-193
380	PD02138	PRECURSOR GLYCOPROTEIN SIGNAL CELL.	PD02138A 27.60 9.057e-09 20-69
381	BL01105	Ribosomal protein L35Ae proteins.	BL01105B 12.95 7.930e-13 43-83
384	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 9.205e-10 10-25 PR00049D 0.00 1.915e-09 9-24
385	BL01115	GTP-binding nuclear protein ran proteins.	BL01115A 10.22 8.909e-13 34-78
385	BL00905	GTP1/OBG family proteins.	BL00905D 15.00 5.313e-09 140-155
385	PR00449	TRANSFORMING PROTEIN P21 RAS SIGNATURE	PR00449C 17.27 3.209e-19 75-98 PR00449A 13.20 1.000e-17 34-56 PR00449D 10.79 3.368e-13 139-153 PR00449B 14.34 8.364e-11 57-74 PR00449E 13.50 8.286e-09 174-197
386	BL00115	Eukaryotic RNA polymerase II heptapeptide repeat proteins.	BL00115Z 3.12 7.977e-10 397-446
386	PR00041	CAMP RESPONSE ELEMENT BINDING (CREB) PROTEIN SIGNATURE	PR00041F 8.53 9.365e-09 256-274
388	PF00646	F-box domain proteins.	PF00646A 14.37 9.036e-10 28-42
389	BL00036	bZIP transcription factors basic domain proteins.	BL00036 9.02 6.294e-12 81-94
389	PR00042	FOS TRANSFORMING PROTEIN SIGNATURE	PR00042C 8.29 8.105e-13 82-99 PR00042D 8.97 9.895e-10 100-122
389	BL00224	Clathrin light chain proteins.	BL00224B 16.94 3.373e-09 70-123
389	PR00043	JUN TRANSCRIPTION FACTOR SIGNATURE	PR00043B 8.73 9.596e-09 81-98
390	PF00622	Domain in SP1a and the RYanodine Receptor.	PF00622B 21.00 2.500e-13 85-107
391	BL00564	Argininosuccinate synthase proteins.	BL00564A 19.93 6.114e-09 7-44
392	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 7.750e-14 230-244 PR00048A 10.52 4.316e-11 202-216
392	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 2.125e-15 205-222 BL00028 16.07 1.391e-12 233-250 BL00028 16.07 3.400e-10 177-194
392	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 3.000e-13 193-206 PD00066 13.92 3.423e-10 221-234
393	PF00622	Domain in SP1a and the RYanodine Receptor.	PF00622B 21.00 1.391e-16 132-154

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
393	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 8.800e-10 761-778 BL00028 16.07 2.029e-09 789-806
393	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 2.800e-09 758-772
394	PR00501	KELCH REPEAT SIGNATURE	PR00501A 8.25 1.409e-09 537-551
394	DM00099	4 kw A55R REDUCTASE TERMINAL DIHYDROPTERIDINE	DM00099B 14.73 4.375e-09 415-425
395	PR00399	SYNAPTOTAGMIN SIGNATURE	PR00399A 9.52 3.133e-19 146-162 PR00399C 12.82 8.200e-17 222-238 PR00399B 14.27 7.750e-16 161-175 PR00399D 14.48 4.000e-14 242-253
395	PR00360	C2 DOMAIN SIGNATURE	PR00360B 13.61 8.269e-13 201-215 PR00360A 14.59 2.800e-12 174-187 PR00360B 13.61 5.217e-12 340-354 PR00360A 14.59 5.207e-10 311-324
395	PF00168	C2 domain proteins.	PF00168C 27.49 5.500e-18 323-349 PF00168B 11.83 2.000e-09 306-317
396	BL01013	Oxysterol-binding protein family proteins.	BL01013A 25.14 7.231e-21 558-156 BL01013B 11.33 1.000e-11 185-196
396	PF00791	Domain present in ZO-1 and Unc5-like netrin receptors.	PF00791B 28.49 3.534e-10 52-107
396	PD00078	REPEAT PROTEIN ANK NUCLEAR ANKYR.	PD00078B 13.14 9.000e-11 173-186 PD00078B 13.14 3.739e-09 78-91 PD00078B 13.14 4.130e-09 45-58
396	PF00023	Ank repeat proteins.	PF00023B 14.20 3.077e-11 48-58 PF00023B 14.20 3.769e-11 176-186 PF00023A 16.03 7.429e-09 85-101
397	PF00023	Ank repeat proteins.	PF00023A 16.03 1.750e-10 55-71
397	PF00791	Domain present in ZO-1 and Unc5-like netrin receptors.	PF00791B 28.49 4.455e-11 55-110 PF00791B 28.49 7.291e-10 88-143
398	BL00422	Granins proteins.	BL00422C 16.18 5.787e-10 134-162
400	PR00450	RECOVERIN FAMILY SIGNATURE	PR00450D 16.58 8.986e-11 161-181
400	BL00479	Phorbol esters / diacylglycerol binding domain proteins.	BL00479B 12.57 4.273e-15 287-303 BL00479A 19.86 2.667e-14 261-284 BL00479B 12.57 1.360e-10 351-367
400	PR00171	CLASS III CYTOCHROME C SIGNATURE	PR00171D 7.30 9.419e-10 334-342
400	BL00018	EF-hand calcium-binding domain proteins.	BL00018 7.41 3.348e-09 223-236
400	PF00781	Diacylglycerol kinase catalytic domain proteins (presumed).	PF00781F 16.43 1.000e-40 600-199 PF00781B 12.07 8.364e-35 454-486 PF00781D 11.11 3.077e-30 532-118 PF00781C 9.69 5.034e-19 506-521 PF00781E 12.45 2.385e-17 124-583 PF00781G 10.09 6.211e-17 678-692 PF00781H 12.20 1.750e-16 770-782 PF00781A 6.42 3.667e-09 354-360
401	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 7.407e-09 325-340
402	DM01117	2 kw TRANSPOSASE WITHIN	DM01117A 11.17 7.750e-09 364-382

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
		TRANSPOSITION VASOTOCIN.	
403	DM01206	CORONAVIRUS NUCLEOCAPSID PROTEIN.	DM01206B 10.69 9.286e-12 724-744 DM01206B 10.69 3.466e-10 726-746 DM01206B 10.69 9.630e-10 722-742 DM01206B 10.69 7.152e-09 718-738 DM01206B 10.69 8.861e-09 728-748
403	BL00048	Protamine P1 proteins.	BL00048 6.39 4.197e-10 722-749 BL00048 6.39 5.500e-10 731-758 BL00048 6.39 6.329e-10 729-756 BL00048 6.39 9.171e-10 730-757 BL00048 6.39 4.038e-09 728-755 BL00048 6.39 8.538e-09 724-751 BL00048 6.39 9.438e-09 716-743
403	PD00289	PROTEIN SH3 DOMAIN REPEAT PRESYNA.	PD00289 9.97 9.690e-09 130-144
404	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 1.353e-27 31-70
404	PD00066	PROTEIN ZINC-FINGER METAL-BINDL.	PD00066 13.92 5.154e-15 274-287 PD00066 13.92 7.600e-14 246-259 PD00066 13.92 8.200e-14 302-315 PD00066 13.92 3.143e-12 218-231 PD00066 13.92 4.000e-12 190-203 PD00066 13.92 2.800e-09 330-343
404	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 7.261e-12 230-247 BL00028 16.07 9.171e-12 342-359 BL00028 16.07 4.300e-10 314-331 BL00028 16.07 7.000e-10 174-191 BL00028 16.07 3.314e-09 202-219 BL00028 16.07 6.400e-09 286-303
404	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 4.214e-13 339-353 PR00048A 10.52 3.209e-12 227-241 PR00048A 10.52 1.947e-11 311-325 PR00048A 10.52 4.522e-10 171-185 PR00048B 6.02 2.895e-09 299-309 PR00048A 10.52 4.600e-09 199-213 PR00048B 6.02 1.000e-08 187-197 PR00048B 6.02 1.000e-08 271-281
406	BL00610	Sodium:neurotransmitter symporter family proteins.	BL00610A 17.73 1.000e-40 68-118 BL00610B 23.65 1.000e-40 132-182 BL00610C 12.94 1.000e-40 225-277 BL00610D 20.97 1.000e-40 291-344 BL00610F 29.02 6.143e-36 540-157 BL00610E 20.34 3.209e-35 448-491 BL00610G 12.89 2.200e-15 173-196
406	PR00176	SODIUM/NEUROTRANSMITTER SYMPORTER SIGNATURE	PR00176C 10.84 6.226e-23 141-168 PR00176A 16.82 1.450e-22 68-90 PR00176F 10.73 8.667e-20 452-472 PR00176B 7.31 7.000e-18 97-117 PR00176D 9.02 1.000e-17 252-270 PR00176E 11.41 2.756e-15 334-355 PR00176H 15.27 7.353e-15 131-590 PR00176G 12.48 5.615e-14 529-112
407	DM00179	w KINASE ALPHA ADHESION T-CELL.	DM00179 13.97 5.304e-09 111-121

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
408	PR00187	DNAJ PROTEIN FAMILY SIGNATURE	PR00187B 13.48 1.800e-16 45-66 PR00187A 12.84 6.700e-12 15-35
408	BL00198	Nt-dnaJ domain proteins.	BL00198B 15.11 9.217e-15 45-66 BL00198A 8.07 2.459e-11 19-36
409	PR00927	ADENINE NUCLEOTIDE TRANSLOCATOR 1 SIGNATURE	PR00927E 14.93 4.136e-11 246-268
409	BL00215	Mitochondrial energy transfer proteins.	BL00215A 15.82 9.735e-14 11-36 BL00215A 15.82 5.787e-11 108-133 BL00215B 10.44 6.211e-11 258-271 BL00215A 15.82 5.018e-09 211-236
409	PR00926	MITOCHONDRIAL CARRIER PROTEIN SIGNATURE	PR00926D 10.53 5.355e-09 19-38
410	PD00066	PROTEIN ZINC-FINGER METAL-BINDL	PD00066 13.92 6.400e-17 411-424 PD00066 13.92 8.200e-17 327-340 PD00066 13.92 5.154e-15 271-284 PD00066 13.92 2.800e-14 215-228 PD00066 13.92 9.000e-13 355-368 PD00066 13.92 6.143e-12 439-452 PD00066 13.92 6.478e-11 187-200 PD00066 13.92 9.217e-11 243-256
410	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 2.588e-14 227-244 BL00028 16.07 6.824e-14 395-412 BL00028 16.07 7.882e-14 171-188 BL00028 16.07 2.350e-13 339-356 BL00028 16.07 7.300e-13 283-300 BL00028 16.07 7.300e-13 367-384 BL00028 16.07 2.565e-12 423-440 BL00028 16.07 7.261e-12 199-216 BL00028 16.07 7.261e-12 311-328 BL00028 16.07 8.435e-12 451-468 BL00028 16.07 2.038e-11 255-272 BL00028 16.07 9.400e-10 143-160
410	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 3.250e-14 280-294 PR00048A 10.52 8.500e-14 336-350 PR00048A 10.52 7.429e-13 252-266 PR00048A 10.52 8.714e-13 448-462 PR00048A 10.52 9.357e-13 392-406 PR00048A 10.52 1.000e-12 168-182 PR00048A 10.52 2.059e-12 420-434 PR00048B 6.02 8.615e-11 408-418 PR00048B 6.02 7.188e-10 268-278 PR00048B 6.02 7.188e-10 380-390 PR00048B 6.02 9.438e-10 296-306 PR00048B 6.02 1.000e-09 324-334 PR00048B 6.02 1.474e-09 352-362 PR00048B 6.02 3.842e-09 212-222 PR00048B 6.02 5.263e-09 436-446
411	BL00018	EF-hand calcium-binding domain proteins.	BL00018 7.41 5.500e-10 63-76
413	PR00014	FIBRONECTIN TYPE III REPEAT SIGNATURE	PR00014C 15.44 4.600e-10 73-92
414	PR00806	VINCULIN SIGNATURE	PR00806A 6.63 1.493e-09 785-796
414	PR00048	C2H2-TYPE ZINC FINGER	PR00048A 10.52 4.240e-09 41-55

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
		SIGNATURE	
414	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 9.546e-11 781-796 PR00049D 0.00 1.205e-10 263-278 PR00049D 0.00 4.356e-09 785-800
414	BL00412	Neuromodulin (GAP-43) proteins.	BL00412D 16.54 4.673e-09 420-471
414	BL00422	Granins proteins.	BL00422C 16.18 6.318e-11 439-467 BL00422C 16.18 9.809e-10 440-468 BL00422C 16.18 6.294e-09 441-469 BL00422C 16.18 6.209e-09 438-466
414	PR00910	LUTEOVIRUS ORF6 PROTEIN SIGNATURE	PR00910A 2.51 8.179e-09 265-278
414	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 4.203e-09 770-803 DM00215 19.43 9.085e-09 245-278
414	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 1.257e-09 44-61 BL00028 16.07 2.543e-09 175-192 BL00028 16.07 6.143e-09 119-136 BL00028 16.07 9.743e-09 147-164
415	PF00622	Receptor.Domain in SPLa and the RYanodine	PF00622B 21.00 1.000e-13 331-353 PF00622C
415	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 3.400e-11 31-40
416	PF00780	Domain found in NIK1-like kinases, mouse citron and yeast ROM.	PF00780B 23.03 5.929e-33 442-485
416	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109B 12.27 5.235e-12 211-230
416	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 5.200e-22 211-242 BL00107B 13.31 9.308e-12 283-299
416	BL00239	Receptor tyrosine kinase class II proteins.	BL00239B 25.15 5.164e-10 145-193
416	BL00915	Phosphatidylinositol 3- and 4-kinases proteins.	BL00915C 22.43 9.357e-10 203-242 *
417	BL00021	Kringle domain proteins.	BL00021B 13.33 1.482e-14 41-59 BL00021D 24.56 2.122e-12 193-235
417	PR00722	CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1) SIGNATURE	PR00722A 12.27 7.517e-14 42-58 PR00722B 12.51 3.143e-10 97-112
417	BL00134	Serine proteases, trypsin family, histidine proteins.	BL00134A 11.96 6.464e-16 41-58 BL00134C 13.45 2.059e-09 221-235
417	BL00495	Apple domain proteins.	BL00495O 13.75 2.440e-09 212-241
417	BL00672	Serine proteases, V8 family, histidine proteins.	BL00672A 9.79 9.520e-09 41-57
417	PR00839	V8 SERINE PROTEASE FAMILY SIGNATURE	PR00839B 11.20 9.753e-09 41-59
418	BL01207	Glypicans proteins.	BL01207B 23.69 9.122e-28 191-237 BL01207A 12.21 1.000e-16 62-78
423	PD02870	RECEPTOR INTERLEUKIN-1 PRECURSOR.	PD02870D 15.74 4.351e-09 693-728
423	DM00179	w KINASE ALPHA ADHESION T-CELL.	DM00179 13.97 5.696e-09 793-803

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
424	BL00203	Vertebrate metallothioneins proteins.	BL00203 13.94 5.041e-09 13-59
425	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 8.141e-18 217-248
425	BL00240	Receptor tyrosine kinase class III proteins.	BL00240E 11.56 6.040e-10 203-241
425	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109B 12.27 5.814e-14 217-236 PR00109A 15.00 1.730e-09 182-196
428	PR00141	PROTEASOME COMPONENT SIGNATURE	PR00141C 11.15 6.333e-12 234-246 PR00141D 12.45 8.615e-12 259-271 PR00141B 11.15 9.561e-12 223-235 PR00141A 11.36 2.050e-11 102-118
428	BL00854	Proteasome B-type subunits proteins.	BL00854A 33.93 1.383e-19 99-145 BL00854C 29.92 5.235e-14 206-235 BL00854D 13.76 2.800e-09 257-267
429	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245A 18.03 9.413e-17 59-81 PR00245C 7.84 7.500e-16 238-254 PR00245E 12.40 2.500e-12 291-306 PR00245B 10.38 9.112e-11 177-192
429	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237E 13.03 7.120e-12 199-223 PR00237C 15.69 1.225e-09 104-127
429	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 9.727e-14 90-130 BL00237D 11.23 1.273e-09 282-299
429	PR00534	MELANOCORTIN RECEPTOR FAMILY SIGNATURE	PR00534A 11.49 6.400e-09 51-64
430	PF00651	BTB (also known as BR-C/Ttk) domain proteins.	PF00651 15.00 1.000e-11 87-100
430	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 4.706e-14 474-491 BL00028 16.07 1.771e-09 502-519
430	PD00066	PROTEIN ZINC-FINGER METAL-BINDL	PD00066 13.92 4.300e-09 490-503
430	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 4.600e-09 499-513
433	BL00086	Cytochrome P450 cysteine heme-iron ligand proteins.	BL00086 20.87 3.209e-23 430-462
433	PR00465	E-CLASS P450 GROUP IV SIGNATURE	PR00465F 13.37 1.360e-11 400-419
433	PR00359	B-CLASS P450 SIGNATURE	PR00359G 11.22 8.071e-10 401-417 PR00359F 24.20 2.180e-09 373-401
433	PR00385	P450 SUPERFAMILY SIGNATURE	PR00385E 12.66 8.800e-11 440-452 PR00385D 13.11 4.429e-10 431-441 PR00385A 14.97 5.865e-09 302-320
433	PR00464	E-CLASS P450 GROUP II SIGNATURE	PR00464G 12.41 9.000e-10 405-421 PR00464D 17.40 1.191e-09 320-338 PR00464E 18.28 6.946e-09 349-370 PR00464H 13.32 7.750e-09 427-441 PR00464C 18.84 9.014e-09 291-320 PR00464I 14.64 9.481e-09 440-464
434	BL00216	Sugar transport proteins.	BL00216B 27.64 7.943e-19 101-151
434	PR00171	SUGAR TRANSPORTER SIGNATURE	PR00171D 12.76 3.593e-11 413-435

Table 3

SEQ ID NO:	Database entry ID	Description	*Results
435	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 8.429e-10 10-25
435	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 4.150e-13 138-593 BL00028 16.07 6.850e-13 1010-1027 BL00028 16.07 6.087e-12 982-999 BL00028 16.07 8.615e-11 846-863 BL00028 16.07 3.100e-10 317-334 BL00028 16.07 7.000e-10 170-187 BL00028 16.07 8.500e-10 289-306 BL00028 16.07 8.800e-10 548-565
435	PD00066	PROTEIN ZINC-FINGER METAL-BINDL	PD00066 13.92 7.600e-14 998-1011 PD00066 13.92 1.000e-11 305-318 PD00066 13.92 8.826e-11 564-577 PD00066 13.92 3.400e-09 862-875
435	PR00456	RIBOSOMAL PROTEIN P2 SIGNATURE	PR00456E 3.06 5.329e-09 177-192 PR00456E 3.06 5.899e-09 140-155
435	BL00999	Streptomyces subtilisin-type inhibitors proteins.	BL00999A 14.95 7.223e-09 461-499
435	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 9.357e-13 573-587 PR00048A 10.52 2.421e-11 1007-1021 PR00048B 6.02 2.125e-10 561-133 PR00048A 10.52 8.043e-10 314-328 PR00048B 6.02 1.000e-09 995-1005 PR00048B 6.02 6.684e-09 302-312 PR00048A 10.52 9.280e-09 167-181
436	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245A 18.03 2.667e-23 100-122 PR00245C 7.84 1.783e-14 232-248 PR00245D 10.47 7.070e-10 268-280
436	PR00237	RHODOPSIN-LIKE GPCR	PR00237C 15.69 8.500e-11 145-168
		SUPERFAMILY SIGNATURE	PR00237G 19.63 6.023e-09 266-293
436	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 2.161e-15 131-171 BL00237D 11.23 8.091e-09 276-293
437	PR00262	IL1/HBGF FAMILY SIGNATURE	PR00262A 28.26 1.000e-08 80-108
438	BL00884	Osteopontin proteins.	BL00884B 12.47 1.000e-40 50-94 BL00884C 22.45 6.187e-39 131-173 BL00884A 11.35 5.846e-32 1-31 BL00884E 11.04 8.364e-23 273-295 BL00884D 8.79 3.323e-18 255-272
438	PR00216	OSTEOPONTIN SIGNATURE	PR00216B 7.89 4.553e-34 37-67 PR00216A 10.94 8.054e-33 2-32 PR00216C 9.63 2.565e-32 67-93 PR00216G 12.39 8.676e-27 238-264 PR00216H 7.41 5.295e-22 273-293 PR00216F 11.79 3.133e-21 164-183 PR00216D 2.74 5.800e-18 104-119 PR00216E 8.44 4.405e-16 132-147

Table 3

* Results include in order: Accession No., subtype, e-value, and amino acid position of the signature in the corresponding polypeptide

Table 4

SEQ ID	Pfam Model	Description	E-value	Score	No: of Pfam Domains	Position of the Domain
1	zf-CCCH	Zinc finger C-x8-C-x5-C-x3-H type	1.8e-05	31.6	1	412-438
1	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	2e-05	21.8	1	14-52
3	EMP24_GP25L	emp24/gp25L/p24 family	4.1e-105	362.6	1	22-235
6	WW	WW domain	1.2e-05	32.2	1	45-75
7	WW	WW domain	1.2e-05	32.2	1	45-75
8	Aa_trans	Transmembrane amino acid transporter protein	9.6e-64	225.2	1	71-451
9	Fe-ADH	Iron-containing alcohol dehydrogenase	9.9e-35	124.5	2	4-205:228-255
10	Fe-ADH	Iron-containing alcohol dehydrogenase	9.9e-35	124.5	2	52-253:276-303
11	Bcl-2	Apoptosis regulator proteins, Bcl-2 family	0.016	-2.1	1	257-356
12	spectrin	Spectrin repeat	1.3e-10	43.6	3	11-87:90-197:200-291
13	Ribosomal_L18ae	Ribosomal L18ae protein family	1.9e-128	440.1	1	6-176
14	Ribosomal_L31e	Ribosomal protein L31e	2.4e-47	170.7	1	72-166
15	zf-CCCH	Zinc finger C-x8-C-x5-C-x3-H type	7.8e-16	66.0	3	342-367:371-396:398-420
16	zf-MYND	MYND finger	1.4e-13	58.5	1	52-90
17	Sterile	Male sterility protein	1.1e-51	185.1	1	254-446
18	MgtE	Divalent cation transporter	8.6e-39	142.3	2	138-274:352-499
19	Rap_GAP	Rap/ran-GAP	2e-124	426.7	1	400-588
19	PDZ	PDZ domain (Also known as DHR or GLGF)	2.4e-06	34.5	1	726-800
20	Rap_GAP	Rap/ran-GAP	2e-124	426.7	1	400-588
20	PDZ	PDZ domain (Also known as DHR or GLGF)	2.4e-06	34.5	1	726-800
22	SCAN	SCAN domain	1.5e-23	91.7	1	165-238
23	RhoGAP	RhoGAP domain	3e-58	206.9	1	497-649
23	FCH	Fes/CIP4 homology domain	1.2e-18	75.4	1	22-121
23	SH3	SH3 domain	2.6e-11	51.0	1	723-777
24	adh_zinc	Zinc-binding dehydrogenases	1.5e-05	-25.4	1	20-336
25	UDPGT	UDP-glucuronosyl and UDP-glucosyl transferas	1.6e-84	294.3	1	26-467
28	Ribosomal_L6e	Ribosomal protein L6e	4.3e-77	269.5	1	109-239
29	Ribosomal_L11	Ribosomal protein L11	4.9e-64	226.2	1	13-144
30	tRNA-synt_1e	tRNA synthetases class I (C)	1.6e-137	470.2	1	64-538
32	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	0.00041	17.6	2	33-72:165-185
34	ras	Ras family	1.4e-77	271.2	1	35-235
34	arf	ADP-ribosylation factor family	9.3e-05	-56.3	1	17-198
36	SET	SET domain	3.2e-05	10.0	1	209-342
36	MORN	MORN repeat	0.006	23.2	3	36-58:59-81:106-128

Table 4

SEQ ID	Pfam Model	Description	E-value	Score	No: of Pfam Domains	Position of the Domain
37	laminin_G	Laminin G domain	1.5e-11	44.7	1	55-174
37	EGF	EGF-like domain	0.0033	24.1	1	202-234
38	Sema	Sema domain	1.7e-127	436.9	1	56-489
38	Plexin_repeat	Plexin repeat	1e-06	35.7	1	507-563
38	ig	Immunoglobulin domain	0.0023	15.9	1	582-639
38	integrin_B	Integrins, beta chain	0.084	6.1	1	513-527
40	filament	Intermediate filament protein	1.6e-138	473.6	1	129-442
41	Keratin_B2	Keratin, high sulfur B2 protein	1.8e-18	74.8	2	2-138:139-240
44	sushi	Sushi domain (SCR repeat)	3.8e-06	33.9	4	1396-1459:1464-1521:1525-1590:1595-1646
45	profilin	Profilin	4.1e-13	51.7	1	10-124
47	ubiquitin	Ubiquitin family	0.00033	20.5	1	31-99
48	BTB	BTB/POZ domain	2.6e-21	84.2	1	80-196
48	Kelch	Kelch motif	2.6e-20	80.9	4	336-382:384-430:432-478:582-635
48	SCP	SCP-like extracellular protein	0.015	13.0	1	1-35
49	serpin	Serpin (serine protease inhibitor)	2.4e-178	605.4	1	59-432
50	T-box	T-box	3.6e-125	429.2	1	140-331
52	7tm_1	7 transmembrane receptor (rhodopsin family)	1.2e-17	58.3	2	132-228:337-344
53	CSD	'Cold-shock' DNA-binding domain	1.8e-16	63.6	1	42-112
53	zf-CCHC	Zinc knuckle	0.00012	28.8	2	137-154:159-176
54	ig	Immunoglobulin domain	2.5e-07	28.7	1	34-109
55	Rap_GAP	Rap/ran-GAP	5e-18	73.3	1	287-466
57	G-gamma	GGL domain	1.8e-11	39.4	2	49-70:109-
58	T-box	T-box	8.9e-114	391.4	1	101-302
59	Gag_p10	Retroviral GAG p10 protein	9.2e-06	23.7	1	82-171
61	60s_ribosomal	60s Acidic ribosomal protein	0.0089	12.0	1	1-22
62	UPAR_LY6	u-PAR/Ly-6 domain	5.4e-05	22.3	1	8-51
63	Ribosomal_L30	Ribosomal protein L30p/L7e	0.00042	18.5	1	65-93
64	filament	Intermediate filament protein	1.1e-78	274.8	2	161-338:339-426
65	Ribosomal_S6	Ribosomal protein S6	0.00082	7.5	1	2-96
66	PDZ	PDZ domain (Also known as DHR or GLGF)	5.1e-09	43.4	1	158-250
67	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	0.005	14.0	1	92-118
68	G-patch	G-patch domain	6.8e-07	36.3	1	26-70
69	Keratin_B2	Keratin, high sulfur B2 protein	0.037	-45.9	1	10-155
83	ig	Immunoglobulin domain	8.5e-09	33.4	2	34-89:119-187
86	zf-C2H2	Zinc finger, C2H2 type	2.2e-71	250.6	17	182-204:210-

Table 4

SEQ ID	Pfam Model	Description	E-value	Score	No: of Pfam Domains	Position of the Domain
						232:237- 260:265- 288:315- 337:343- 365:369- 392:653- 675:681- 704:709- 733:741- 764:791- 814:820- 842:848- 870:877- 899:905- 928:952-975
87	ig	Immunoglobulin domain	2.7e-35	118.7	6	36-121:162- 249:292- 375:422- 517:564- 657:704-795
88	MAP1_LC3	Microtubule associated protein 1A/1B, light	9.4e-79	275.0	1	118-221
89	WD40	WD domain, G-beta repeat	1.6e-12	55.1	4	173-215:221- 263:269- 305:1103- 1140
90	FKBP	FKBP-type peptidyl-prolyl cis-trans isomerases	1.2e-59	198.9	1	66-160
92	RPEL	RPEL repeat	6.5e-18	73.0	2	513-538:551- 576
93	transket_pyr	Transketolase, pyridine binding domain	4.6e-65	229.6	1	568-773
93	E1_dehydrog	Dehydrogenase E1 component	8.7e-23	89.1	1	193-504
95	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	8.7e-09	32.7	1	595-635
97	ig	Immunoglobulin domain	1.8e-20	71.0	3	31-88:127- 185:222-278
98	ig	Immunoglobulin domain	1.8e-20	71.0	3	24-81:120- 178:215-271
99	Patched	Patched family	6.2e-06	-369.1	1	66-935
102	zf-C2H2	Zinc finger, C2H2 type	2.3e-94	326.9	12	209-231:237- 259:265- 287:293- 315:321- 343:349- 371:377- 399:405- 427:433- 455:461- 483:489- 511:594-616

Table 4

SEQ ID	Pfam Model	Description	E-value	Score	No: of Pfam Domains	Position of the Domain
102	KRAB	KRAB box	3.7e-37	136.9	1	15-77
103	zf-C2H2	Zinc finger, C2H2 type	1.2e-55	198.2	9	172-195:271-293:299-321:327-349:355-377:383-405:411-433:439-461:467-489
103	KRAB	KRAB box	3e-46	167.1	1	8-70
107	zf-CCHC	Zinc knuckle	2.4e-16	67.8	3	913-930:1293-1310:1358-1375
107	NTP_transf_2	Nucleotidyltransferase domain	4.4e-11	50.3	1	972-1065
108	zf-C2H2	Zinc finger, C2H2 type	1.6e-42	154.7	5	283:289-311:317-339:345-367:373-395
109	myosin_head	Myosin head (motor domain)	0	1267.5	1	26-697
109	IQ	IQ calmodulin-binding motif	1.2e-17	72.1	4	714-734:737-757:760-780:789-809
110	pkinase	Protein kinase domain	1.2e-96	334.5	1	20-271
111	WD40	WD domain, G-beta repeat	1.8e-49	177.8	8	161-197:218-253:258-294:300-335:341-377:383-428:434-470:476-511
112	SNF2_N	SNF2 and others N-terminal domain	4.2e-78	272.9	1	1-264
112	helicase_C	Helicase conserved C-terminal domain	1.2e-24	95.4	1	326-410
113	DUF15	Domain of unknown function DUF15	0.00064	-60.4	1	132-384
114	DSPc	Dual specificity phosphatase, catalytic	0.0004	-2.9	1	141-295
114	Y_phosphatase	Protein-tyrosine phosphatase	0.0037	-26.9	1	128-295
115	Ulp1_C	Ulp1 protease family, C-terminal catalytic d	2.8e-52	187.1	1	394-587
117	Rhodanese	Rhodanese-like domain	1e-05	32.4	1	160-260
119	ABC1	ABC1 family	1.7e-40	147.9	1	318-434
122	proteasome	Proteasome A-type and B-type	7.4e-43	155.8	1	39-146
124	Ribosomal_L9	Ribosomal protein L9	3.1e-05	-3.4	1	94-240
125	RIO1	RIO1/ZK632.3/MJ0444 family	7.8e-80	278.6	1	193-387
128	abhydrolase	alpha/beta hydrolase fold	4.5e-20	80.1	1	121-364
129	TPR	TPR Domain	4.8e-27	103.3	7	355-388:473-

Table 4

SEQ ID	Pfam Model	Description	E-value	Score	No: of Pfam Domains	Position of the Domain
						506:507-540:654-687:688-721:722-755:756-789
130	HMG14_17	HMG14 and HMG17	1.9e-15	64.7	1	2-73
131	bZIP	bZIP transcription	8.3e-19	71.7	1	288-352
132	rrm	RNA recognition motif.	1.9e-31	117.9	3	432-502:546-616:858-929
133	AMP-binding	AMP-binding enzyme	7.1e-117	401.7	1	142-580
138	tubulin	Tubulin/FtsZ family	2.1e-151	516.4	1	1-223
141	laminin_EGF	Laminin EGF-like (Domains III and V)	7.6e-12	52.8	4	252-297:300-348:1342-1391:1469-1530
141	Kelch	Kelch motif	1.6e-05	31.8	4	654-702:760-811:873-918:929-990
141	integrin_B	Integrins, beta chain	0.0061	9.4	3	44-59:100-117:1019-1028
141	EGF	EGF-like domain	0.092	19.3	8	167-203:207-235:297-331:496-533:538-569:1271-1308:1312-1338:1478-1508
142	RUN	RUN domain	8e-44	159.0	1	31-163
142	FYVE	FYVE zinc finger	2.3e-29	109.1	1	529-593
143	zf-C2H2	Zinc finger, C2H2 type	1.7e-33	124.7	5	442-464:505-527:533-555:561-583:589-611
143	BTB	BTB/POZ domain	1.6e-22	88.2	1	30-143
144	mito_carr	Mitochondrial carrier protein	3.6e-61	216.6	3	10-158:160-250:254-354
146	DAGKc	Diacylglycerol kinase catalytic domain	0.00015	26.0	1	157-303
147	Exonuclease	Exonuclease	1.6e-41	151.4	1	228-384
147	rrm	RNA recognition motif.	9.5e-08	39.2	2	507-574:602-674
151	WH2	WH2 motif	6.5e-20	79.6	3	1194-1214:1234-1254:1322-1342
154	DHDPS	Dihydrodipicolinate synthetase family	9.1e-21	82.4	1	3-270
156	PseudoU_synth_1	tRNA pseudouridine synthase	1e-30	115.4	1	111-322
157	pkinase	Protein kinase domain	2.3e-59	210.6	1	216-512
158	ubiquitin	Ubiquitin family	2.4e-05	24.6	1	3-79

Table 4

SEQ ID	Pfam Model	Description	E-value	Score	No: of Pfam Domains	Position of the Domain
160	IF-2B	Initiation factor 2 subunit family	1.7e-98	340.7	1	157-475
161	Beach	Beige/BEACH domain	1.1e-224	759.8	1	1470-1747
161	WD40	WD domain, G-beta repeat	2.9e-08	40.9	5	1848-1882:1888-1928:1947-1983:2030-2064:2071-2107
164	DnaJ	DnaJ domain	1.9e-16	68.1	1	125-189
165	Anti_proliferat	BTG1 family	7.4e-85	295.3	1	11-164
166	sugar_tr	Sugar (and other) transporter	1.2e-78	274.7	1	34-548
167	sugar_tr	Sugar (and other) transporter	7e-52	185.8	1	34-480
168	zf-C2H2	Zinc finger, C2H2 type	1.7e-93	324.0	13	222-244:250-272:278-300:306-328:334-356:362-384:390-412:418-440:446-468:474-496:502-524:530-552:558-580
168	KRAB	KRAB box	1.8e-35	131.2	1	57-119
169	GBP	Guanylate-binding protein, N-terminal domain	1e-191	636.2	1	1-275
169	GBP_C	Guanylate-binding protein, C-terminal domain	6.6e-162	551.3	1	277-573
170	cyclin	Cyclin, N-terminal domain	0.0022	9.3	1	48-192
171	TPR	TPR Domain	9.7e-43	155.4	6	133-166:167-200:201-234:282-315:316-349:350-383
173	RhoGEF	RhoGEF domain	3.3e-40	147.0	1	166-345
173	PH	PH domain	6.5e-14	54.5	1	378-483
173	SH3	SH3 domain	1.1e-10	48.9	1	72-126
174	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	0.00011	19.4	1	18-55
174	GBP_C	Guanylate-binding protein, C-terminal domain	0.016	12.1	1	86-114
175	Peptidase_M22	Glycoprotease family	2.3e-73	257.2	1	1-324
177	TBC	TBC domain	4.7e-08	10.1	1	57-268
178	transmembrane4	Tetraspanin family	1.6e-78	259.2	1	16-261
179	CH	Calponin homology (CH) domain	1.2e-25	98.6	1	24-133
179	calponin	Calponin family repeat	1.7e-14	51.8	1	174-199
182	AP_endonucleas1	AP endonuclease family 1	2.6e-17	59.4	2	1-36:50-135
184	Bacterial_PQQ	PQQ enzyme repeat	9.3e-05	29.2	2	52-89:534-571

Table 4

SEQ ID	Pfam Model	Description	E-value	Score	No: of Pfam Domains	Position of the Domain
185	DEAD	DEAD/DEAH box helicase	1.6e-60	194.3	1	216-420
185	helicase_C	Helicase conserved C-terminal domain	5.9e-25	96.3	1	454-540
186	zf-C2H2	Zinc finger, C2H2 type	3.2e-24	93.9	6	106-128:134-156:162-184:195-218:477-499:505-529
187	sugar_tr	Sugar (and other) transporter	0.0014	-90.1	1	272-672
188	tRNA_int_endo	tRNA intron endonuclease, catalytic C-t	0.0025	-7.7	1	73-159
189	WSC	WSC domain	1e-35	132.1	1	175-254
189	Sulfotransfer	Sulfotransferase protein	4e-34	126.8	1	356-586
191	pkinase	Protein kinase domain	5.1e-75	262.6	1	148-421
191	PDZ	PDZ domain (Also known as DHR or GLGF)	1.3e-05	32.1	1	740-827
193	globin	Globin	1.9e-26	96.6	1	3-78
195	WD40	WD domain, G-beta repeat	6.7e-14	59.6	4	64-108:116-153:158-194:288-323
197	BRO1	BRO1-like domain	0.0042	-29.4	1	9-161
198	F_actin_cap_B	F-actin capping protein, beta subunit	1.7e-224	759.2	1	1-269
199	ank	Ank repeat	1e-66	235.0	8	40-73:82-114:115-147:148-180:181-212:213-246:481-526:527-559
203	PDZ	PDZ domain (Also known as DHR or GLGF)	4.2e-07	37.0	1	211-293
204	SAM	SAM domain (Sterile alpha motif)	1.2e-11	52.1	1	5-70
205	SAM	SAM domain (Sterile alpha motif)	1.2e-11	52.1	1	5-70
206	zf-UBR1	Putative zinc finger in N-recognin	4.7e-25	96.7	1	978-1046
207	ABC_tran	ABC transporter	2.4e-112	386.6	2	467-647:1536-1717
209	zf-C2H2	Zinc finger, C2H2 type	0.00035	27.3	1	200-225
210	UCH-2	Ubiquitin carboxyl-terminal hydrolase family	1.5e-19	78.4	1	385-454
211	IMP4	Domain of unknown function	2.2e-33	124.3	1	144-297
213	zf-C2H2	Zinc finger, C2H2 type	2.9e-08	40.9	3	12-37:173-198:208-230
214	LysM	LysM domain	2.1e-11	51.3	1	73-116
215	ank	Ank repeat	1.1e-05	32.3	2	834-867:879-912
215	TIG	IPT/TIG domain	0.009	22.6	1	642-723
217	pyr_redox	Pyridine nucleotide-	1.7e-71	251.0	1	196-470

Table 4

SEQ ID	Pfam Model	Description	E-value	Score	No: of Pfam Domains	Position of the Domain
		disulphide oxidoreducta				
217	Rieske	Rieske [2Fe-2S] domain	6.2e-20	79.6	1	68-168
218	PDZ	PDZ domain (Also known as DHR or GLGF)	8.5e-19	75.9	1	642-728
219	pkinese	Protein kinase domain	8.1e-67	235.4	1	26-204
220	dsrm	Double-stranded RNA binding motif	0.095	7.5	1	100-172
221	PHD	PHD-finger	5.4e-05	29.6	1	147-203
222	L27	L27 domain	6.5e-16	66.3	1	13-68
222	SAM	SAM domain (Sterile alpha motif)	7.2e-10	46.2	2	1051-1117:1166-1230
223	TRM	N2,N2-dimethylguanosine tRNA methyltransfera	7.3e-22	86.1	1	227-693
224	LIM	LIM domain	5.3e-06	33.4	2	124-180:183-243
225	ig	Immunoglobulin domain	1.1e-07	29.8	1	55-144
227	F-box	F-box domain	1.3e-05	32.1	1	11-59
229	Glucosamine_iso	Glucosamine-6-phosphate isomerases/6-	2.7e-158	539.3	1	15-250
231	PTN_MK	PTN/MK heparin-binding protein family	3.6e-44	160.2	1	51-148
236	ion_trans	Ion transport protein	1.6e-22	88.3	1	174-393
238	GNS1_SUR4	GNS1/SUR4 family	5.2e-46	166.3	1	10-265
240	ubiquitin	Ubiquitin family	2.7e-05	24.4	1	10-89
241	PIP5K	Phosphatidylinositol-4-phosphate 5-Kinase	1.5e-155	530.2	1	124-420
242	cadherin	Cadherin domain	0	1298.9	19	1-75:89-180:194-290:355-434:448-549:563-652:671-774:788-881:896-988:1002-1092:1106-1192:1206-1295:1309-1379:1393-1489:1503-1594:1608-1699:1713-1808:1814-1910:1922-2016
244	fn3	Fibronectin type III domain	1.2e-31	118.6	4	58-140:152-238:249-333:345-426
245	UQ_con	Ubiquitin-conjugating enzyme	1.4e-16	68.5	1	93-250
246	LRR	Leucine Rich Repeat	1.7e-14	61.6	6	51-75:76-

Table 4

SEQ ID	Pfam Model	Description	E-value	Score	No: of Pfam Domains	Position of the Domain
						99:155-178:181-203:204-226:227-251
247	lipocalin	Lipocalin / cytosolic fatty-acid binding pr	1.2e-28	102.8	1	164-294
248	Ribosomal_S2	Ribosomal protein S2	2.9e-11	43.7	1	33-80
249	tubulin	Tubulin/FtsZ family	8.5e-163	554.2	1	1-277
250	tubulin	Tubulin/FtsZ family	2.4e-212	718.8	1	1-351
251	ATP-synt_ab	ATP synthase alpha/beta family, nucleot	1.2e-75	264.8	1	138-346
251	ATP-synt_ab_C	ATP synthase alpha/beta chain, C termin	2.7e-38	140.6	1	348-456
251	ATP-synt_ab_N	ATP synthase alpha/beta family, beta-ba	5.4e-19	76.5	1	67-135
252	ATP-synt_ab	ATP synthase alpha/beta family, nucleot	1.3e-70	248.0	1	138-344
252	ATP-synt_ab_N	ATP synthase alpha/beta family, beta-ba	5.4e-19	76.5	1	67-135
253	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	5e-12	43.2	1	39-79
254	G-patch	G-patch domain	1.3e-08	42.1	1	410-456
255	CH	Calponin homology (CH) domain	1.6e-11	51.7	1	24-134
256	RF-1	Peptidyl-tRNA hydrolase domain	5.9e-66	232.5	1	225-338
257	RF-1	Peptidyl-tRNA hydrolase domain	5.9e-66	232.5	1	189-302
258	OTU	OTU-like cysteine protease	4.4e-18	73.5	1	189-304
259	thioered	Thioredoxin	2e-09	35.7	2	119-165:662-695
260	thyroglobulin_1	Thyroglobulin type-1 repeat	3.1e-34	127.2	2	95-158:227-292
260	kazal	Kazal-type serine protease inhibitor	9.3e-07	35.9	1	43-87
262	DnaJ	DnaJ domain	4.1e-15	63.6	1	277-338
263	WD40	WD domain, G-beta repeat	4e-21	83.6	5	3-42:49-86:97-133:142-178:184-220
265	DUF6	Integral membrane protein DUF6	0.083	9.1	2	81-316:338-470
266	Ribosomal_L31e	Ribosomal protein L31e	1.7e-61	217.7	1	15-109
268	F5_F8_type_C	F5/8 type C domain	2.4e-65	230.5	1	42-196
268	Zn_carbOpept	Zinc carboxypeptidase	3.5e-50	180.1	2	224-341:400-600
270	BTB	BTB/POZ domain	7.7e-18	72.7	1	8-119
270	zf-C2H2	Zinc finger, C2H2 type	4.2e-13	57.0	4	254-276:363-385:390-412:448-468
271	Glycos_transf_1	Glycosyl transferases group 1	0.027	12.8	1	291-385
272	HEAT	HEAT repeat	2.2e-07	38.0	3	237-275:276-

Table 4

SEQ ID	Pfam Model	Description	E-value	Score	No: of Pfam Domains	Position of the Domain
						315:674-712
273	HEAT	HEAT repeat	2.2e-07	38.0	3	237-275:276-315:640-678
275	SPRY	SPRY domain	2.6e-34	127.4	1	390-515
275	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	1e-16	58.5	1	29-69
277	BTB	BTB/POZ domain	6e-27	103.0	1	36-149
277	Kelch	Kelch motif	9.7e-21	82.3	4	331-390:392-441:443-493:540-586
278	zf-C2H2	Zinc finger, C2H2 type	4.1e-116	399.2	14	193-215:221-243:249-271:277-299:305-327:333-355:361-383:389-411:417-439:445-467:473-495:501-523:529-551:557-579
229	SCAN	SCAN domain	2.4e-52	187.3	1	36-132
229	zf-C2H2	Zinc finger, C2H2 type	2.4e-51	184.0	7	348-370:375-397:403-425:431-453:459-480:486-508:514-537
231	Zip	ZIP Zinc transporter	6.6e-20	79.6	1	1-146
282	NTP_transf_2	Nucleotidyltransferase domain	8.5e-13	55.9	1	67-174
286	zf-C2H2	Zinc finger, C2H2 type	2.8e-93	323.3	12	118-140:146-168:174-196:202-224:230-252:258-280:286-308:314-336:342-364:370-392:398-420:426-448
286	KRAB	KRAB box	3.6e-38	140.2	1	8-70
287	zf-C2H2	Zinc finger, C2H2 type	5.3e-124	425.4	17	183-205:211-233:239-261:267-289:295-317:323-345:351-373:379-

Table 4

SEQ ID	Pfam Model	Description	E-value	Score	No: of Pfam Domains	Position of the Domain
						401:407-429:435-457:463-485:491-513:519-541:547-569:575-597:603-625:631-653
289	DiHfolate_red	Dihydrofolate reductase	7.4e-77	268.8	1	4-185
291	PDZ	PDZ domain (Also known as DHR or GLGF)	7.4e-17	69.4	1	5-84
293	PH	PH domain	1.4e-08	35.5	1	44-147
294	adh_short	short chain dehydrogenase	3.9e-29	110.2	1	36-284
297	PKD	PKD domain	9.9e-09	42.4	2	663-753:756-839
297	BNR	BNR repeat	3.2e-06	34.1	5	115-126:156-167:351-362:428-439:470-481
300	HMG_box	HMG (high mobility group) box	5.4e-05	20.0	1	245-304
301	ig	Immunoglobulin domain	0.05	11.6	1	629-688
302	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	5e-12	43.2	1	39-79
303	START	START domain	0.015	4.1	1	1790-1994
304	integrase	Integrase DNA binding domain	7.2e-06	32.9	1	51-96
305	myosin_head	Myosin head (motor domain)	7.6e-279	939.7	2	11-668:689-733
306	zf-C2H2	Zinc finger, C2H2 type	8.5e-54	192.1	7	66-88:94-116:122-144:150-172:178-200:280-303:317-339
307	ig	Immunoglobulin domain	0.00023	19.1	2	35-104:136-194
309	ras	Ras family	0.00079	-93.3	1	38-176
310	ig	Immunoglobulin domain	2.1e-06	25.7	1	37-112
311	EF1BD	EF-1 guanine nucleotide exchange domain	4.7e-56	199.6	1	139-225
312	BTB	BTB/POZ domain	8.4e-25	95.8	1	51-164
313	zf-C2H2	Zinc finger, C2H2 type	7.7e-59	208.9	9	118-140:197-219:281-303:309-331:337-359:365-387:393-415:421-443:449-471
313	KRAB	KRAB box	1.4e-17	71.8	1	41-99

Table 4

SEQ ID	Pfam Model	Description	E-value	Score	No: of Pfam Domains	Position of the Domain
314	Hydrolase	haloacid dehalogenase-like hydrolase	0.045	8.2	1	213-671
315	cNMP_binding	Cyclic nucleotide-binding domain	4e-26	100.2	1	387-475
315	ion_trans	Ion transport protein	3.8e-19	77.0	1	69-290
316	Peptidase_S26	Signal peptidase I	2.8e-16	56.3	2	38-98:117-139
317	zf-C2H2	Zinc finger, C2H2 type	4.3e-56	199.8	9	156-178:184-206:212-234:240-262:268-290:296-318:324-346:352-374:378-400
317	KRAB	KRAB box	6.7e-16	66.3	1	11-73
319	UPF0073	Uncharacterised protein family	1.8e-09	27.9	1	33-276
320	EGF	EGF-like domain	4.7e-08	40.2	1	26-59
321	lectin_c	Lectin C-type domain	8.6e-15	62.6	1	268-374
325	MAM	MAM domain	1.3e-52	188.2	1	338-503
325	ig	Immunoglobulin domain	1.9e-15	54.8	3	41-101:138-202:346-420
327	MAM	MAM domain	5.3e-180	611.4	4	26-169:170-329:342-498:509-666
328	Sema	Sema domain	1.5e-211	716.2	1	56-491
329	zf-C2H2	Zinc finger, C2H2 type	1.5e-84	294.3	13	170-192:198-220:226-248:254-276:282-304:310-332:338-360:366-388:394-416:422-444:450-472:478-500:506-528
331	PAP2	PAP2 superfamily	8e-22	85.9	1	160-314
332	LRR	Leucine Rich Repeat	3.4e-36	133.7	11	58-81:82-105:106-129:130-153:154-177:178-201:202-225:250-273:274-297:298-321:322-345
332	ig	Immunoglobulin domain	2.5e-08	31.9	1	425-485
332	LRRNT	Leucine rich repeat N-	2.5e-05	31.1	1	27-56

Table 4

SEQ ID	Pfam Model	Description	E-value	Score	No: of Pfam Domains	Position of the Domain
		terminal domain				
332	LRRCT	Leucine rich repeat C-terminal domain	0.0029	24.3	1	355-408
333	AdoHcyase	S-adenosyl-L-homocysteine hydrolase	1.5e-280	945.4	1	214-640
334	TBC	TBC domain	9.4e-38	138.9	1	89-302
341	WD40	WD domain, G-beta repeat	0.00094	25.9	2	2-32:109-146
342	ABC1	ABC1 family	0.051	-29.9	1	3-50
344	globin	Globin	3e-45	162.2	1	1-141
345	globin	Globin	7.5e-39	139.9	2	1-31:68-179
347	F-box	F-box domain	1.5e-07	38.5	1	24-72
348	HLH	Helix-loop-helix DNA-binding domain	2e-08	41.4	1	83-137
349	KRAB	KRAB box	2.7e-39	144.0	1	4-66
350	UCH-2	Ubiquitin carboxyl-terminal hydrolase family	1.7e-19	78.2	1	645-705
350	UCH-1	Ubiquitin carboxyl-terminal hydrolases famil	9.1e-15	62.5	1	363-394
350	zf-UBP	Zn-finger in ubiquitin-hydrolases and other	0.00069	18.9	1	236-306
351	NUDIX	MutT-like domain	8.2e-12	52.7	1	50-200
352	IBR	IBR domain	1.6e-12	55.0	1	101-166
353	IBR	IBR domain	1.6e-12	55.0	1	66-131
354	SCP	SCP-like extracellular protein	1.4e-34	128.3	1	56-208
356	mito_carr	Mitochondrial carrier protein	9.7e-78	271.7	3	10-125:127-220:232-321
358	UCH-1	Ubiquitin carboxyl-terminal hydrolases famil	5.1e-15	63.3	1	323-354
358	zf-UBP	Zn-finger in ubiquitin-hydrolases and other	0.00049	19.4	1	195-264
360	Phage_lysozyme	Phage lysozyme	0.0014	23.4	1	94-184
362	Ribosomal_S2	Ribosomal protein S2	3.3e-08	32.9	1	20-62
364	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	5.3e-09	33.4	1	291-329
365	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	0.0096	13.1	1	109-148
367	TPR	TPR Domain	0.043	20.4	1	1-28
370	zf-C2H2	Zinc finger, C2H2 type	5.3e-109	375.5	14	127-149:155-177:183-205:211-233:239-261:267-289:295-317:323-345:351-373:379-401:407-429:435-457:463-485:491-513
370	SCAN	SCAN domain	4.2e-38	140.0	1	27-122
371	arf	ADP-ribosylation factor	4.9e-39	143.1	1	6-184

Table 4

SEQ ID	Pfam Model	Description	E-value	Score	No: of Pfam Domains	Position of the Domain
		family				
371	ras	Ras family	7.2e-06	-70.1	1	22-186
372	BNR	BNR repeat	0.031	20.9	3	171-182:244-255:295-306
373	zf-C2H2	Zinc finger, C2H2 type	8.3e-25	95.8	5	142-162:171-198:204-228:234-258:264-288
376	rrm	RNA recognition motif.	0.00019	28.2	1	112-163
377	rrm	RNA recognition motif.	2.2e-19	77.9	1	112-183
380	vwc	von Willebrand factor type C domain	1.6e-31	118.2	3	22-76:79-134:137-192
381	Ribosomal_L35Ae	Ribosomal protein L35Ae	0.00013	7.0	1	1-79
385	ras	Ras family	3.9e-63	223.2	1	35-229
385	arf	ADP-ribosylation factor family	1.7e-05	-46.9	1	18-202
388	F-box	F-box domain	1.5e-05	31.9	2	23-70:99-146
390	SPRY	SPRY domain	6.2e-10	46.4	1	101-239
391	tRNA_Me_trans	tRNA methyl transferase	1.9e-19	50.9	1	5-185
392	zf-C2H2	Zinc finger, C2H2 type	4e-17	70.3	3	175-197:203-225:231-253
393	SCAN	SCAN domain	3.1e-39	143.8	1	389-484
393	SPRY	SPRY domain	1.8e-19	78.1	1	148-273
393	zf-C2H2	Zinc finger, C2H2 type	4e-09	43.7	2	759-781:787-809
393	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	0.0032	14.7	1	11-52
394	Kelch	Kelch motif	4e-53	189.9	5	329-375:377-431:433-479:481-525:527-572
394	BTB	BTB/POZ domain	6.1e-26	99.6	1	30-144
395	C2	C2 domain	2.2e-80	280.4	2	159-251:296-384
396	ank	Ank repeat	5.6e-33	123.0	4	47-79:80-112:140-174:175-207
396	PH	PH domain	8.9e-05	22.0	1	236-334
397	ank	Ank repeat	1.7e-26	101.4	4	17-49:50-82:83-115:116-148
398	Nucleoplasmin	Nucleoplasmin	3.6e-29	110.4	1	13-209
400	DAGKa	Diacylglycerol kinase accessory domain	1.9e-124	426.8	1	598-778
400	DAGKc	Diacylglycerol kinase catalytic domain	7.1e-67	235.6	1	454-578
400	DAG_PE-bind	Phorbol esters/diacylglycerol binding dom	2.9e-23	90.7	2	261-310:326-374
400	efhand	EF hand	2.4e-12	54.4	2	169-197:214-242
403	PDZ	PDZ domain (Also known as	7.7e-46	165.7	3	86-166:210-

Table 4

SEQ ID	Pfam Model	Description	E-value	Score	No. of Pfam Domains	Position of the Domain
		DHR or GLGF)				291:821-907
404	zf-C2H2	Zinc finger, C2H2 type	2.6e-48	173.9	7	172-194:200-222:228-250:256-278:284-306:312-331:340-362
405	K_tetra	K+ channel tetramerisation domain	2.6e-23	90.9	1	51-146
406	SNF	Sodium:neurotransmitter symporter family	0	1268.7	1	60-657
407	ig	Immunoglobulin domain	1.1e-06	26.5	1	53-120
408	DnaJ	DnaJ domain	2.3e-27	104.3	1	4-68
408	DnaJ_C	DnaJ C terminal region	3.1e-08	38.1	1	192-314
409	mito_carr	Mitochondrial carrier protein	1.4e-57	204.7	3	5-100:102-201:205-302
410	zf-C2H2	Zinc finger, C2H2 type	5.2e-97	335.7	12	141-163:169-191:197-219:225-247:253-275:281-303:309-331:337-359:365-387:393-415:421-443:449-473
411	S_100	S-100/ICaBP type calcium binding domain	9.7e-13	55.8	1	5-48
411	efhand	EF hand	0.0012	25.6	1	54-82
413	fn3	Fibronectin type III domain	8.6e-14	59.3	2	22-107:119-196
413	PHD	PHD-finger	9.6e-05	27.2	1	285-341
414	zf-C2H2	Zinc finger, C2H2 type	2.3e-27	104.4	6	42-64:117-139:145-167:173-196:534-556:573-595
415	SPRY	SPRY domain	3.9e-18	73.7	1	347-467
415	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	4.4e-14	49.9	1	16-56
415	zf-B_box	B-box zinc finger	9e-07	35.9	1	92-133
416	pkinase	Protein kinase domain	1.2e-54	195.0	1	97-317
417	trypsin	Trypsin	4.6e-38	122.5	1	41-234
418	Glypican	Glypican	5.7e-131	448.5	1	3-244
419	Keratin_B2	Keratin, high sulfur B2 protein	0.0013	-23.4	1	37-159
420	Dynein_heavy	Dynein heavy chain	0	1432.3	1	309-1019
421	zf-C2H2	Zinc finger, C2H2 type	0.00039	27.2	3	75-99:203-227:266-290
422	ig	Immunoglobulin domain	0.00074	17.5	1	34-107
423	fn3	Fibronectin type III domain	6e-08	39.8	1	443-531

Table 4

SEQ ID	Pfam Model	Description	E-value	Score	No: of Pfam Domains	Position of the Domain
424	Keratin_B2	Keratin, high sulfur B2 protein	0.0023	-27.1	2	5-150:152-251
425	pkinase	Protein kinase domain	2.3e-55	197.3	1	69-390
426	ig	Immunoglobulin domain	4.1e-09	34.4	1	35-112
427	Galactosyl_T	Galactosyltransferase	2.6e-35	130.8	1	158-349
428	proteasome	Proteasome A-type and B-type	5.5e-28	106.4	1	96-238
429	7tm_1	7 transmembrane receptor (rhodopsin family)	3.4e-38	123.5	1	41-290
430	BTB	BTB/POZ domain	8.1e-23	89.2	1	58-173
430	zf-C2H2	Zinc finger, C2H2 type	4.3e-07	37.0	2	472-494:500-523
433	p450	Cytochrome P450	6.4e-175	594.5	1	33-493
434	sugar_tr	Sugar (and other) transporter	2.6e-64	227.1	1	10-512
435	zf-C2H2	Zinc finger, C2H2 type	1.8e-52	187.8	9	287-309:315-337:546-568:574-596:606-628:844-866:872-894:980-1002:1008-1030
436	7tm_1	7 transmembrane receptor (rhodopsin family)	2.2e-40	130.4	2	82-221:229-284
437	FGF	Fibroblast growth factor	4.6e-14	51.6	1	48-129
438	Osteopontin	Osteopontin	3.7e-181	615.2	1	1-294

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
1	1bor		10	60	5.2e-09	-0.58	0.01		TRANSCRIPTION FACTOR PML; CHAIN: NULL;	TRANSCRIPTION REGULATION PROTO-ONCOGENE, NUCLEAR BODIES (PODS), LEUKEMIA, 2 TRANSCRIPTION REGULATION
1	1dvp	A	5	72	0.0061	-0.23	0.76		HEPATOCYTE GROWTH FACTOR-REGULATED TYROSINE CHAIN: A;	TRANSFERASE HRS; HRS, VHS, FYVE, ZINC FINGER, SUPERHELIX
1	1g25	A	10	60	1.3e-06	-0.44	0.62		CDK-ACTIVATING KINASE ASSEMBLY FACTOR MAT1; CHAIN: A;	METAL BINDING PROTEIN RING FINGER PROTEIN MAT1; RING FINGER (C3HC4)
6	1pin	A	48	86	7.8e-06	-0.45	0.23		PEPTIDYL-PROLYL CIS-TRANS ISOMERASE; CHAIN: A; ALA-PRO DIPEPTIDE; CHAIN: B;	COMPLEX (ISOMERASE/DIPEPTIDE) PIN1; PEPTIDYL-PROLYL CIS-TRANS ISOMERASE, ROTAMASE, 2 COMPLEX (ISOMERASE/DIPEPTIDE) CONECT
11	1bxi	A	223	362	5.1e-38	-0.07	0.01		BCL-XL; CHAIN: A; BAK PEPTIDE; CHAIN: B;	COMPLEX (APOPTOSIS/PEPTIDE) APOPTOSIS, ALTERNATIVE SPLICING, COMPLEX (APOPTOSIS/PEPTIDE)
11	1fl6	A	208	362	1.7e-36	-0.28	0.10		APOPTOSIS REGULATOR BAX, MEMBRANE ISOFORM ALPHA; CHAIN: A;	APOPTOSIS HELICAL PROTEIN
11	1maz		223	362	3.4e-40	-0.03	0.03		BCL-XL; CHAIN: NULL;	APOPTOSIS APOPTOSIS REGULATOR BCL-X; APOPTOSIS, PROGRAMMED CELL DEATH, BCL-2 FAMILY
12	1cum	A	94	297	1e-21			81.85	ALPHA SPECTRIN; CHAIN:	STRUCTURAL PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									A, B, C;	TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
12	1quu	A	77	297	3.4e-11			72.83	HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
14	1ffk	U	75	142	2.6e-26	-0.90	0.00		23S RRNA; CHAIN: O; 5S RRNA; CHAIN: 9; RIBOSOMAL PROTEIN L2; CHAIN: A; RIBOSOMAL PROTEIN L3; CHAIN: B; RIBOSOMAL PROTEIN L4; CHAIN: C; RIBOSOMAL PROTEIN L5; CHAIN: D; RIBOSOMAL PROTEIN L7AE; CHAIN: E; RIBOSOMAL PROTEIN L10E; CHAIN: F; RIBOSOMAL PROTEIN L13; CHAIN: G; RIBOSOMAL PROTEIN L14; CHAIN: H; RIBOSOMAL PROTEIN L15E; CHAIN: I; RIBOSOMAL PROTEIN L15; CHAIN: J; RIBOSOMAL PROTEIN L18; CHAIN: K; RIBOSOMAL PROTEIN L18E; CHAIN: L; RIBOSOMAL PROTEIN L19; CHAIN: M; RIBOSOMAL PROTEIN L21E; CHAIN: N;	RIBOSOME 50S RIBOSOMAL PROTEIN L2P, HMAL2, HL4; 50S RIBOSOMAL PROTEIN L3P, HMAL3, HL1; 50S RIBOSOMAL PROTEIN L4E, HMAL4, HL6; 50S RIBOSOMAL PROTEIN L5P, HMAL5, HL13; 30S RIBOSOMAL PROTEIN HS6; 50S RIBOSOMAL PROTEIN L13P, HMAL13; 50S RIBOSOMAL PROTEIN L14P, HMAL14, HL27; 50S RIBOSOMAL PROTEIN L15P, HMAL15, HL9; 50S RIBOSOMAL PROTEIN L18P, HMAL18, HL12; 50S RIBOSOMAL PROTEIN L18E, HL29, L19; 50S RIBOSOMAL PROTEIN L19E, HMAL19, HL24; 50S RIBOSOMAL PROTEIN L21E, HL31; 50S RIBOSOMAL PROTEIN L22P, HMAL22, HL23; 50S RIBOSOMAL PROTEIN L23P, HMAL23, HL25, L21; 50S RIBOSOMAL PROTEIN L24P, HMAL24, HL16, HL15; 50S

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									RIBOSOMAL PROTEIN L22; CHAIN: O; RIBOSOMAL PROTEIN L23; CHAIN: P; RIBOSOMAL PROTEIN L24; CHAIN: Q; RIBOSOMAL PROTEIN L24E; CHAIN: R; RIBOSOMAL PROTEIN L29; CHAIN: S; RIBOSOMAL PROTEIN L30; CHAIN: T; RIBOSOMAL PROTEIN L31E; CHAIN: U; RIBOSOMAL PROTEIN L32E; CHAIN: V; RIBOSOMAL PROTEIN L37AB; CHAIN: W; RIBOSOMAL PROTEIN L37E; CHAIN: X; RIBOSOMAL PROTEIN L39E; CHAIN: Y; RIBOSOMAL PROTEIN L44E; CHAIN: Z; RIBOSOMAL PROTEIN L6; CHAIN: 1;	RIBOSOMAL PROTEIN L24E, HL21/HL22; 50S RIBOSOMAL PROTEIN L29P, HMAL29, HL33; 50S RIBOSOMAL PROTEIN L30P, HMAL30, HL20, HL16; 50S RIBOSOMAL PROTEIN L31E, L34, HL30; 50S RIBOSOMAL PROTEIN L32E, HL5; 50S RIBOSOMAL PROTEIN L37E, L35E; 50S RIBOSOMAL PROTEINS L39E, HL39E, HL46E; 50S RIBOSOMAL PROTEIN L44E, LA, HLA; 50S RIBOSOMAL PROTEIN L6P, HMAL6, HL10 RIBOSOME ASSEMBLY, RNA-RNA, PROTEIN-RNA, PROTEIN- PROTEIN
14	1ffk	U	78	152	5.1e-21	-0.40	0.12		23S RRNA; CHAIN: 0; 5S RRNA; CHAIN: 9; RIBOSOMAL PROTEIN L2; CHAIN: A; RIBOSOMAL PROTEIN L3; CHAIN: B; RIBOSOMAL PROTEIN L4; CHAIN: C; RIBOSOMAL PROTEIN L5; CHAIN: D; RIBOSOMAL PROTEIN L7AE; CHAIN: E; RIBOSOMAL PROTEIN L10E; CHAIN: F; RIBOSOMAL PROTEIN	RIBOSOME 50S RIBOSOMAL PROTEIN L2P, HMAL2, HL4; 50S RIBOSOMAL PROTEIN L3P, HMAL3, HL1; 50S RIBOSOMAL PROTEIN L4E, HMAL4, HL6; 50S RIBOSOMAL PROTEIN L5P, HMAL5, HL13; 30S RIBOSOMAL PROTEIN HS6; 50S RIBOSOMAL PROTEIN L13P, HMAL13; 50S RIBOSOMAL PROTEIN L14P, HMAL14, HL27; 50S

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									L13; CHAIN: G; RIBOSOMAL PROTEIN L14; CHAIN: H; RIBOSOMAL PROTEIN L15E; CHAIN: I; RIBOSOMAL PROTEIN L15; CHAIN: J; RIBOSOMAL PROTEIN L18; CHAIN: K; RIBOSOMAL PROTEIN L18E; CHAIN: L; RIBOSOMAL PROTEIN L19; CHAIN: M; RIBOSOMAL PROTEIN L21E; CHAIN: N; RIBOSOMAL PROTEIN L22; CHAIN: O; RIBOSOMAL PROTEIN L23; CHAIN: P; RIBOSOMAL PROTEIN L24; CHAIN: Q; RIBOSOMAL PROTEIN L24E; CHAIN: R; RIBOSOMAL PROTEIN L29; CHAIN: S; RIBOSOMAL PROTEIN L30; CHAIN: T; RIBOSOMAL PROTEIN L31E; CHAIN: U; RIBOSOMAL PROTEIN L32E; CHAIN: V; RIBOSOMAL PROTEIN L37AE; CHAIN: W; RIBOSOMAL PROTEIN L37E; CHAIN: X; RIBOSOMAL PROTEIN L39E; CHAIN: Y; RIBOSOMAL PROTEIN L44E; CHAIN: Z;	RIBOSOMAL PROTEIN L15P, HMA15, HL9; 50S RIBOSOMAL PROTEIN L18P, HMA18, HL12; 50S RIBOSOMAL PROTEIN L18E, HL29, L19; 50S RIBOSOMAL PROTEIN L19E, HMA19, HL24; 50S RIBOSOMAL PROTEIN L21E, HL31; 50S RIBOSOMAL PROTEIN L22P, HMA122, HL23; 50S RIBOSOMAL PROTEIN L23P, HMA123, HL25, L21; 50S RIBOSOMAL PROTEIN L24P, HMA124, HL16, HL15; 50S RIBOSOMAL PROTEIN L24E, HL21/HL22; 50S RIBOSOMAL PROTEIN L29P, HMA129, HL33; 50S RIBOSOMAL PROTEIN L30P, HMA130, HL20, HL16; 50S RIBOSOMAL PROTEIN L31E, L34, HL30; 50S RIBOSOMAL PROTEIN L32E, HL5; 50S RIBOSOMAL PROTEIN L37E, L35E; 50S RIBOSOMAL PROTEINS L39E, HL39E, HL46E; 50S RIBOSOMAL PROTEIN L44E, LA, HLA; 50S RIBOSOMAL PROTEIN L6P, HMA16, HL10 RIBOSOME ASSEMBLY, RNA-RNA, PROTEIN-RNA, PROTEIN- PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									RIBOSOMAL PROTEIN L6; CHAIN: 1;	
16	1elr	A	322	420	3.4e-11	0.04	-0.15		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
16	1fch	A	174	450	5.1e-17	-0.24	0.04		PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR; CHAIN: A, B; PTS1-CONTAINING PEPTIDE; CHAIN: C, D;	SIGNALING PROTEIN PEROXISOMORE RECEPTOR 1, PTS1-BP, PEROXIN-5, PTS1 PROTEIN-PEPTIDE COMPLEX, TETRATRICOPEPTIDE REPEAT, TPR, 2 HELICAL REPEAT
17	1bxk	A	11	361	1e-72			55.95	DTDP-GLUCOSE 4-6-DEHYDRATASE; CHAIN: A, B;	LYASE EPIMERASE, DEHYDRATASE, LYASE
17	1udb		11	363	5.1e-63			72.32	UDP-GALACTOSE-4-EPIMERASE; CHAIN: NULL;	ISOMERASE EPIMERASE; UDP-GALACTOSE, EPIMERASE, ISOMERASE
25	1f0k	A	208	410	5.1e-20	0.19	0.59		UDP-N-ACETYLGLUCOSAMINE-N-ACETYLMURAMYL-CHAIN: A, B;	TRANSFERASE MURG; ROSSMANN FOLD
30	1a8h		71	563	1.7e-57			103.94	METHIONYL-TRNA SYNTHETASE; CHAIN: NULL;	AMINOACYL-TRNA SYNTHETASE METRS; AMINOACYL-TRNA SYNTHETASE, ROSSMANN FOLD
30	1gln		73	537	8.5e-40			74.49	GLUTAMYL-TRNA SYNTHETASE; 1GLN 4 CHAIN: NULL 1GLN 5	AMINOACYL-TRNA SYNTHASE
32	1chc		33	76	5.1e-11	0.18	0.95		VIRUS EQUINE HERPES	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
32	1fbv	A	31	81	2.6e-13	0.12	0.33		VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	LIGASE CBL, UBCH7, ZAP-70, E2, UBIQUITIN, E3, PHOSPHORYLATION, 2 TYROSINE KINASE, UBIQUITINATION, PROTEIN DEGRADATION,
32	1fbv	A	33	76	3.4e-10	0.73	0.93		SIGNAL TRANSDUCTION PROTEIN CBL; CHAIN: A; ZAP-70 PEPTIDE; CHAIN: B; UBIQUITIN-CONJUGATING ENZYME E12-18 KDA UBCH7; CHAIN: C;	LIGASE CBL, UBCH7, ZAP-70, E2, UBIQUITIN, E3, PHOSPHORYLATION, 2 TYROSINE KINASE, UBIQUITINATION, PROTEIN DEGRADATION,
32	1g25	A	32	78	2.6e-13	0.05	0.25		CDK-ACTIVATING KINASE ASSEMBLY FACTOR MAT1; CHAIN: A;	METAL BINDING PROTEIN RING FINGER PROTEIN MAT1; RING FINGER (C3HC4)
32	1g25	A	33	82	3.4e-06	-0.26	0.51		CDK-ACTIVATING KINASE ASSEMBLY FACTOR MAT1; CHAIN: A;	METAL BINDING PROTEIN RING FINGER PROTEIN MAT1; RING FINGER (C3HC4)
32	1rmd		29	121	3.4e-20	0.25	1.00		RAG1; CHAIN: NULL;	DNA-BINDING PROTEIN V(D)J RECOMBINATION ACTIVATING PROTEIN 1; RAG1, V(D)J RECOMBINATION, ANTIBODY, MAD, RING FINGER, 2 ZINC BINUCLEAR CLUSTER, ZINC FINGER, DNA-BINDING PROTEIN
32	1rmd		9	121	3.4e-20			60.57	RAG1; CHAIN: NULL;	DNA-BINDING PROTEIN V(D)J RECOMBINATION ACTIVATING PROTEIN 1; RAG1, V(D)J RECOMBINATION,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										ANTIBODY, MAD, RING FINGER, 2 ZINC BINUCLEAR CLUSTER, ZINC FINGER, DNA-BINDING PROTEIN
33	1cun	A	2	137	2.6e-14	0.36	0.04		ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
33	1dnl	B	7	166	6.5e-18	-0.03	0.27		SYNTAXIN BINDING PROTEIN 1; CHAIN: A; SYNTAXIN 1A; CHAIN: B; SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS S NSEC1; PROTEIN-PROTEIN COMPLEX, MULTI-SUBUNIT
33	1ez3	A	7	120	2.6e-14	0.09	-0.09			ENDOCYTOSIS/EXOCYTOSIS S SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
33	1qqe	A	12	76	6.5e-05	0.19	0.04		VESICULAR TRANSPORT PROTEIN SEC17; CHAIN: A;	PROTEIN TRANSPORT HELIX-TURN-HELIX TPR-LIKE REPEAT, PROTEIN TRANSPORT
34	1cly	A	32	197	1.4e-64			123.90	RAS-RELATED PROTEIN RAP-1A; CHAIN: A; PROTO-ONCOGENE SERINE/THREONINE PROTEIN KINASE CHAIN: B;	SIGNALING PROTEIN GTP-BINDING PROTEINS, PROTEIN-PROTEIN COMPLEX, EFFECTORS
34	1kao		32	198	6.8e-58			117.80	RAP2A; CHAIN: NULL;	GTP-BINDING PROTEIN GTP-BINDING PROTEIN, SMALL G PROTEIN, RAP2, GDP, RAS
34	1zbd	A	30	202	6.8e-68			152.38	RAB-3A; CHAIN: A; RABPHILIN-3A; CHAIN: B;	COMPLEX (GTP-BINDING/EFFECTOR) RAS-RELATED PROTEIN RAB3A; COMPLEX (GTP-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
34	3rab	A	29	198	1.4e-68			169.98	RAB3A; CHAIN: A;	BINDING/EFFECTOR, G PROTEIN, EFFECTOR, RABCD, 2 SYNAPTIC EXOCYTOSIS, RAB PROTEIN, RAB3A, RABPHILIN
40	1sig		119	432	5.2e-05			79.28	RNA POLYMERASE PRIMARY SIGMA FACTOR; CHAIN: NULL;	HYDROLASE G PROTEIN, VESICULAR TRAFFICKING, GTP HYDROLYSIS, RAB 2 PROTEIN, NEUROTRANSMITTER RELEASE, HYDROLASE
41	1aut	L	128	233	2.6e-08			50.24	ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE, PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
41	1bx7		55	105	3.9e-11	0.03	-0.17		HIRUSTASIN; CHAIN: NULL;	ANTI-COAGULANT ANTI-COAGULANT, PEPTIDIC INHIBITORS, CONFORMATIONAL 2 FLEXIBILITY, SERINE PROTEASE INHIBITOR
41	1eis	A	115	194	2.6e-09	0.04	0.65		AGGLUTININ ISOLECTIN VIAAGGLUTININ ISOLECTIN V; CHAIN: A;	SUGAR BINDING PROTEIN UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPERANTIGEN
41	1eis	A	41	134	3.9e-10	-0.21	0.09		AGGLUTININ ISOLECTIN	SUGAR BINDING PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
41	1eis	A	90	169	1.3e-08	-0.10	0.41		VI/AGGLUTININ ISOLECTIN V; CHAIN: A;	UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPERANTIGEN
41	1en2	A	113	194	1.3e-10	-0.12	0.62		AGGLUTININ ISOLECTIN VI/AGGLUTININ ISOLECTIN V; CHAIN: A;	SUGAR BINDING PROTEIN UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPERANTIGEN
41	1en2	A	6	74	9.1e-11	-0.71	0.11		AGGLUTININ ISOLECTIN I/AGGLUTININ ISOLECTIN V/CHAIN: A;	SUGAR BINDING PROTEIN UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPERANTIGEN, SACCHARIDE BINDING
41	1en2	A	85	169	1.3e-10	-0.05	0.03		AGGLUTININ ISOLECTIN I/AGGLUTININ ISOLECTIN V/CHAIN: A;	SUGAR BINDING PROTEIN UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPERANTIGEN, SACCHARIDE BINDING
41	1klo		65	221	6.5e-22			70.40	LAMININ; CHAIN: NULL;	GLYCOPROTEIN
41	1kst		93	155	1.3e-05	-0.45	0.00		AGGREGATION INHIBITOR, GP ANTAGONIST KISTRIN (NMR, 8 STRUCTURES) IKST 3	GLYCOPROTEIN
41	1pfx	L	17	148	3.9e-16			59.55	FACTOR IXA; CHAIN: C; L.; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
41	1vmo	A	28	196	5.2e-38	0.04	-0.19		MEMBRANE PROTEIN VITELLINE MEMBRANE OUTER LAYER PROTEIN I IVMO 3	
44	1ckl	A	1464	1597	2.6e-11	0.31	-0.08		CD46; CHAIN: A, B, C, D, E, F;	GLYCOPROTEIN MEMBRANE COFACTOR PROTEIN (MCP); VIRUS RECEPTOR, COMPLEMENT COFACTOR, SHORT CONSENSUS REPEAT, 2 SCR, MEASLES VIRUS, GLYCOPROTEIN
44	1ckl	A	1525	1647	1.3e-17	-0.09	0.94		CD46; CHAIN: A, B, C, D, E, F;	GLYCOPROTEIN MEMBRANE COFACTOR PROTEIN (MCP); VIRUS RECEPTOR, COMPLEMENT COFACTOR, SHORT CONSENSUS REPEAT, 2 SCR, MEASLES VIRUS, GLYCOPROTEIN
44	1dva	L	836	924	5.1e-12	0.14	0.04		DES-GLA FACTOR VIA (HEAVY CHAIN); CHAIN: H, I, DES-GLA FACTOR VIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE- ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN- PEPTIDE COMPLEX
44	1e5g	A	1395	1523	2.6e-09	0.40	0.27		COMPLEMENT CONTROL PROTEIN; CHAIN: A;	COMPLEMENT INHIBITOR VCP, SP35; COMPLEMENT, NMR, MODULES, PROTEIN STRUCTURE, VACCINIA VIRUS
44	1e5g	A	1467	1596	2.6e-12	0.29	-0.02		COMPLEMENT CONTROL PROTEIN; CHAIN: A;	COMPLEMENT INHIBITOR VCP, SP35; COMPLEMENT, NMR, MODULES, PROTEIN STRUCTURE, VACCINIA VIRUS

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
44	1e5g	A	1525	1633	3.9e-10	-0.33	0.17		COMPLEMENT CONTROL PROTEIN; CHAIN: A;	COMPLEMENT INHIBITOR VCP, SP35; COMPLEMENT, NMR, MODULES, PROTEIN STRUCTURE, VACCINIA VIRUS
44	1enn		1054	1127	1.7e-09	0.17	-0.17		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
44	1enn		73	137	1.7e-09	0.35	-0.20		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
44	1fak	L	836	924	5.1e-12	0.05	-0.13		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, COFACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
44	1fsb		108	143	5.1e-09	0.25	-0.20		P-SELECTIN; CHAIN: NULL;	CELL ADHESION PROTEIN H, EGF-LIKE DOMAIN, CELL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
44	1hfi		1522	1596	2.6e-09	0.04	0.05		GLYCOPROTEIN FACTOR H, 15TH C-MODULE PAIR (NMR, MINIMIZED AVERAGED 1HFI 1 STRUCTURE) 1HFI 4 1HFI 5	ADHESION PROTEIN, TRANSMEMBRANE, 2 GLYCOPROTEIN
44	1kdo		335	526	3.4e-10	0.00	-0.18		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
44	1qfk	L	841	924	3.4e-11	0.14	0.03		COAGULATION FACTOR VIIA (LIGHT CHAIN); CHAIN: L; COAGULATION FACTOR VIIA (HEAVY CHAIN); CHAIN: H; TRIPEPTIDYL INHIBITOR; CHAIN: C;	SERINE PROTEASE FVIIA; FVIIA; BLOOD COAGULATION, SERINE PROTEASE
44	1qu0	A	283	434	7.8e-12	-0.00	-0.12		LAMININ ALPHA2 CHAIN; CHAIN: A, B, C, D;	METAL BINDING PROTEIN BETA SANDWICH, CALCIUM-BINDING PROTEIN, METAL BINDING 2 PROTEIN
44	1qub	A	1456	1780	3.9e-19			113.83	HUMAN BETA2-GLYCOPROTEIN I; CHAIN: A;	MEMBRANE ADHESION SHORT CONSENSUS REPEAT, SUSHI, COMPLEMENT CONTROL PROTEIN, 2 N-GLYCOSYLATION, MULTI-DOMAIN, MEMBRANE ADHESION
44	1qub	A	1461	1709	1.7e-12	0.16	0.22		HUMAN BETA2-GLYCOPROTEIN I; CHAIN: A;	MEMBRANE ADHESION SHORT CONSENSUS REPEAT, SUSHI, COMPLEMENT CONTROL PROTEIN, 2 N-GLYCOSYLATION, MULTI-DOMAIN, MEMBRANE ADHESION

Table 5

SEQ ID NO;	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
44	1vvc		1330	1450	1.2e-09	0.06	-0.18		VACCINIA VIRUS COMPLEMENT CONTROL PROTEIN; CHAIN: NULL;	COMPLEMENT INHIBITOR SP35, VCP, VACCINIA VIRUS SP35; COMPLEMENT INHIBITOR, COMPLEMENT MODULE, SCR, SUSHI DOMAIN, 2 MODULE PAIR
44	1vvc		1524	1623	3.9e-13	-0.14	0.01		VACCINIA VIRUS COMPLEMENT CONTROL PROTEIN; CHAIN: NULL;	COMPLEMENT INHIBITOR SP35, VCP, VACCINIA VIRUS SP35; COMPLEMENT INHIBITOR, COMPLEMENT MODULE, SCR, SUSHI DOMAIN, 2 MODULE PAIR
44	9wga	A	1072	1252	1.7e-14	0.09	-0.17		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
44	9wga	A	1123	1284	5.1e-12	0.09	-0.19		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
45	1acf		14	128	8.5e-38	1.04	1.00		PROTEIN BINDING PROFILIN I 1ACF 3	
45	1acf		3	129	8.5e-38			69.76	PROTEIN BINDING PROFILIN I 1ACF 3	
45	1cqa		18	128	1.4e-46	0.91	1.00		PROTEIN BINDING PROFILIN I 1ACF 4A	PROTEIN BINDING ACTIN-BINDING PROTEIN, ALLERGEN
45	1dlj	A	16	125	3.4e-22	0.70	0.94		PROFILIN II; CHAIN: A, B, C, D;	CONTRACTILE PROTEIN ACIDIC PROFILIN ISOFORM
45	1f2k	A	14	128	1.2e-39	0.75	1.00		PROFILIN II; CHAIN: A, B;	ACTIN-BINDING PROTEIN, POLY-L-2 PROLINE BINDING PROTEIN, CONTRACTILE PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
45	1f1l		16	125	1.4e-21	0.15	0.94		PROFILIN; CHAIN: NULL;	SEVEN-STRANDED INCOMPLETE ANTIPARALLEL UP-AND-DOWN BETA 2 BARREL, ACTIN-BINDING PROTEIN, POLY-L-PROLINE BINDING 3 PROTEIN, PIP2 BINDING PROTEIN
45	1g5u	A	16	128	1e-45	0.66	0.99		PROFILIN; CHAIN: A, B;	PROTEIN BINDING ACETYLATION, ACTIN-BINDING PROTEIN, MULTIGENE FAMILY
45	1pne		16	125	1.2e-20	0.41	0.98		ACTIN BINDING PROTEIN PROFILIN 1PNE 3	ALLERGEN ALLERGEN, ACTIN-BINDING PROTEIN
45	1ypr	A	14	125	3.4e-38	0.54	1.00		PROFILIN; CHAIN: A, B;	ACTIN-BINDING PROTEIN, ACTIN-BINDING PROTEIN, PROFILIN, CYTOSKELETON
45	1ypr	A	5	129	3.4e-38			70.17	PROFILIN; CHAIN: A, B;	ACTIN-BINDING PROTEIN, ACTIN-BINDING PROTEIN, PROFILIN, CYTOSKELETON
45	3nul		17	126	3.4e-44	0.72	1.00		PROFILIN I; CHAIN: NULL;	ACTIN-BINDING PROTEIN, PROFILIN, CYTOSKELETON
47	1a5r		9	105	6.8e-05	0.28	0.00		SUMO-1; CHAIN: NULL;	TARGETING PROTEIN PIC1, GMP1, UBL1, SENTRIN; SUMO-1, POST-TRANSLATIONAL PROTEIN MODIFICATION, 2 UBIQUITIN-LIKE PROTEINS, TARGETING PROTEIN
47	1b0l	A	31	102	1.2e-23	1.02	1.00		UBIQUITIN-LIKE PROTEIN 7, RUB1; CHAIN: A;	SIGNALING PROTEIN RUB1, UBIQUITIN-LIKE PROTEIN, ARABIDOPSIS, SIGNALING PROTEIN
47	1c3t	A	31	105	1e-31	0.86	1.00		1D8 UBIQUITIN; CHAIN: A;	DE NOVO PROTEIN, PROTEIN DESIGN,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
47	1udd	A	31	103	3.4e-23	1.10	1.00		UBIQUITIN-LIKE PROTEIN NEDD8; CHAIN: A, B, C, D;	SIGNALING PROTEIN NEDD8; NEDD8, NEDD8-LIKE, UBIQUITIN-LIKE, PROTEOLYSIS, SIGNALING 2 PROTEIN
47	1tbe	B	31	101	6.8e-32	1.14	1.00		UBIQUITIN	
47	1ubi		31	105	6.8e-34	1.28	1.00		TETRAUBIQUITIN 1TBE 3 CHROMOSOMAL PROTEIN	
47	1ud7	A	31	105	1.7e-32	0.94	1.00		UBIQUITIN CORE	UBIQUITIN UBIQUITIN, DESIGNED CORE MUTANT
47	1vcb	A	31	89	5.1e-05	0.36	0.90		MUTANT 1D7; CHAIN: A; ELONGIN B; CHAIN: A, D, G, J; ELONGIN C; CHAIN: B, E, H, K; VHL; CHAIN: C, F, I, L;	TRANSCRIPTION TUMOR SUPPRESSOR, CANCER, UBIQUITIN, BETA SANDWICH, 2 TRANSCRIPTION, TRANSCRIPTIONAL ELONGATION
48	1buo	A	70	193	6.8e-22	0.10	0.70		PROMYELOCYTIC LEUKEMIA ZINC FINGER PROTEIN PLZF; CHAIN: A;	GENE REGULATION POZ DOMAIN; PROTEIN-PROTEIN INTERACTION DOMAIN, TRANSCRIPTIONAL 2 REPRESSOR, ZINC-FINGER PROTEIN, X-RAY CRYSTALLOGRAPHY, 3 PROTEIN STRUCTURE, PROMYELOCYTIC LEUKEMIA, GENE REGULATION
48	1gof		339	655	1e-12	0.52	1.00		OXIDOREDUCTASE(OXYGEN(A)) GALACTOSE OXIDASE (E.C.1.1.3.9) (PH 4.5) IGOF 3	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
49	1qlp	A	59	435	0			300.18	ALPHA-1-ANTITRYPSIN; CHAIN: A;	SERINE PROTEASE INHIBITOR ALPHA-1-PROTEINASE INHIBITOR, ALPHA-1-ANTIPROTEINASE; SERINE PROTEASE INHIBITOR, SERPIN, GLYCOPROTEIN, SIGNAL, 2 POLYMORPHISM, EMPHYSEMA, DISEASE MUTATION, ACUTE PHASE
49	1qlp	A	61	434	0	0.76	1.00		ALPHA-1-ANTITRYPSIN; CHAIN: A;	SERINE PROTEASE INHIBITOR ALPHA-1-PROTEINASE INHIBITOR, ALPHA-1-ANTIPROTEINASE; SERINE PROTEASE INHIBITOR, SERPIN, GLYCOPROTEIN, SIGNAL, 2 POLYMORPHISM, EMPHYSEMA, DISEASE MUTATION, ACUTE PHASE
49	1qnn	A	62	434	0	0.72	1.00		ALPHA-1-ANTITRYPSIN; CHAIN: A;	SERINE PROTEASE INHIBITOR, PARTIAL LOOP 2 INSERTION, LOOP-SHEET POLYMERIZATION, EMPHYSEMA, DISEASE 3 MUTATION, ACUTE PHASE PROTEIN, CONFORMATIONAL DISEASE
49	2ach	B	405	433	1.3e-07	-0.73	0.06		PROTEINASE INHIBITOR ALPHA1 ANTICHYMOTRYPSIN 2ACH 3	
50	1xbr	A	138	335	6.8e-44			178.20	T PROTEIN; CHAIN: A, B; DNA; CHAIN: C, D;	COMPLEX (TRANSCRIPTION FACTOR/DNA) COMPLEX (TRANSCRIPTION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
50	1xbr	A	140	330	6.8e-44	0.68	1.00		T PROTEIN; CHAIN: A, B; DNA; CHAIN: C, D;	FACTOR/DNA), TRANSCRIPTION FACTOR, 2 DNA-BINDING PROTEIN COMPLEX (TRANSCRIPTION FACTOR/DNA) COMPLEX (TRANSCRIPTION FACTOR/DNA), TRANSCRIPTION FACTOR, 2 DNA-BINDING PROTEIN
51	1cs6	A	14	71	1.2e-05	-0.71	0.06		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
51	1cs6	A	21	106	1.3e-05	-0.31	0.37		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
51	1cvs	D	17	71	6.5e-05	-0.49	0.28		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN- LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
51	2fcb	A	21	98	5.2e-05	-0.30	0.18		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
53	1alt	A	125	178	3.4e-17	0.07	0.27		NUCLEOCAPSID PROTEIN; CHAIN: A; SL3 STEM- LOOP RNA; CHAIN: B;	COMPLEX (NUCLEOCAPSID PROTEIN/RNA) NUCLEOCAPSID PROTEIN, COMPLEX (NUCLEOCAPSID PROTEIN/RNA), 2 STEM- LOOP RNA
53	1alt	A	151	181	3.4e-08	0.05	-0.13		NUCLEOCAPSID PROTEIN; CHAIN: A; SL3 STEM- LOOP RNA; CHAIN: B;	COMPLEX (NUCLEOCAPSID PROTEIN/RNA) NUCLEOCAPSID PROTEIN, COMPLEX (NUCLEOCAPSID PROTEIN/RNA), 2 STEM- LOOP RNA
53	1aaf		125	178	6.8e-17	0.28	0.12		NUCLEOCAPSID PROTEIN	

Table 5

53	1bj6	A	138	178	1e-14	-0.14	0.15		HIV-1 NUCLEOCAPSID PROTEIN (MN ISOLATE) (NMR, 20 STRUCTURES) 1AAP 3	COMPLEX (NUCLEOCAPSID PROTEIN/DNA) (12-53) NCPT; COMPLEX (NUCLEOCAPSID PROTEIN/DNA), NUCLEIC ACID, 2 RETROVIRUS, VIRUS MORPHOGENESIS, ZINC FINGER
53	1c9o	A	40	107	1.7e-10	0.29	0.69		COLD-SHOCK PROTEIN; CHAIN: A, B;	ANTIFREEZE PROTEIN CSPB BETA BARREL, HOMODIMER
53	1csp		42	107	1.7e-10	0.20	0.40		TRANSCRIPTION REGULATION MAJOR COLD SHOCK PROTEIN (CSPB) 1CSP 3	
53	1csp		42	112	1.3e-19	0.64	0.87		TRANSCRIPTION REGULATION MAJOR COLD SHOCK PROTEIN (CSPB) 1CSP 3	
53	1mjc		39	106	1e-10	0.21	0.83		TRANSCRIPTION REGULATION MAJOR COLD SHOCK PROTEIN 7.4 (CSPA (CS 7.4)) OF 1MJC 3 (ESCHERICHIA COLI) 1MJC 4	
53	1nc8		133	159	6.8e-05	-0.02	0.06		NUCLEOCAPSID PROTEIN; CHAIN: NULL;	NUCLEOCAPSID PROTEIN, NUCLEOCAPSID PROTEIN, HIV-2, RNA RECOGNITION, ZINC FINGER
53	1nc8		154	181	5.1e-08	0.26	-0.06		NUCLEOCAPSID PROTEIN; CHAIN: NULL;	NUCLEOCAPSID PROTEIN, NUCLEOCAPSID PROTEIN, HIV-2, RNA RECOGNITION, ZINC FINGER
54	1ao7	D	21	134	3.4e-34		65.08		HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA-A2 HEAVY CHAIN; CLASS I MHC, T-CELL RECEPTOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
54	1ao7	E	21	203	5.1e-28			54.73	RECEPTOR ALPHA; CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	VIRAL PEPTIDE, 2 COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR
54	1ay1	L	19	148	6.8e-35	0.24	0.30		HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL RECEPTOR ALPHA; CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA-A2 HEAVY CHAIN: CLASS I MHC, T-CELL RECEPTOR, VIRAL PEPTIDE, 2 COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR
54	1ay1	L	19	204	6.8e-35			51.86	TP7 FAB; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, ANTIBODY, FAB, ENZYME INHIBITOR, PCR, 2 HOT START
54	1b88	A	20	130	1.7e-38			67.06	T CELL RECEPTOR V-ALPHA DOMAIN; CHAIN: A, B;	T CELL RECEPTOR TCR; T CELL RECEPTOR, MHC CLASS I, HUMAN IMMUNODEFICIENCY VIRUS, 2 MOLECULAR RECOGNITION
54	1b88	A	20	142	1.7e-38	0.30	0.92		T CELL RECEPTOR V-ALPHA DOMAIN; CHAIN: A, B;	T CELL RECEPTOR TCR; T CELL RECEPTOR, MHC CLASS I, HUMAN IMMUNODEFICIENCY VIRUS, 2 MOLECULAR RECOGNITION
54	1bd2	D	20	196	1.7e-37	0.26	1.00		HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL RECEPTOR ALPHA;	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA-A2 HEAVY CHAIN: COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
54	1bd2	D	20	204	1.7e-37			70.90	CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA A2 HEAVY CHAIN; COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR)
54	1bwm	A	19	129	3.4e-40	0.13	0.88		ALPHA-BETA T CELL RECEPTOR (TCR) (D10); CHAIN: A;	IMMUNE SYSTEM IMMUNOGLOBULIN, IMMUNORECEPTOR, IMMUNE SYSTEM
54	1d9k	A	21	129	6.8e-39	0.51	1.00		T-CELL RECEPTOR D10 (ALPHA CHAIN); CHAIN: A, E; T-CELL RECEPTOR D10 (BETA CHAIN); CHAIN: B, F; MHC I-AK A CHAIN (ALPHA CHAIN); CHAIN: C, G; MHC I-AK B CHAIN (BETA CHAIN); CHAIN: D, H; CONALBUMIN PEPTIDE; CHAIN: P, Q;	IMMUNE SYSTEM MHC I-AK; MHC I-AK; T-CELL RECEPTOR, MHC CLASS II, D10, I-AK
54	1dee	B	19	165	6.8e-39	0.12	0.52		IGM RF 2A2; CHAIN: A, C, E; IGM RF 2A2; CHAIN: B, D, F; IMMUNOGLOBULIN G BINDING PROTEIN A; CHAIN: G, H;	IMMUNE SYSTEM FAB-IBP COMPLEX CRYSTAL STRUCTURE 2.7A RESOLUTION BINDING 2 OUTSIDE THE ANTIGEN COMBINING SITE SUPERANTIGEN FAB VH3 3 SPECIFICITY
54	1gc1	L	19	204	6.8e-33			50.38	ENVELOPE PROTEIN GP120; CHAIN: G; CD4; CHAIN: C; ANTIBODY 17B; CHAIN: L, H;	COMPLEX (HIV ENVELOPE PROTEIN/CD4/FAB) COMPLEX (HIV ENVELOPE PROTEIN/CD4/FAB), HIV-1 EXTERIOR 2 ENVELOPE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
54	1igt	B	19	168	3.4e-35	0.01	0.45		IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	GPI20, T-CELL SURFACE GLYCOPROTEIN CD4, 3 ANTIGEN-BINDING FRAGMENT OF HUMAN IMMUNOGLOBULIN 17B, 4 GLYCOSYLATED PROTEIN
54	1jrh	L	19	192	3.4e-27			50.28	ANTIBODY A6; CHAIN: L, H; INTERFERON-GAMMA RECEPTOR ALPHA CHAIN; CHAIN: L;	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN COMPLEX (ANTIBODY/ANTIGEN) CYTOKINE RECEPTOR, COMPLEX (ANTIBODY/ANTIGEN), 2 TRANSMEMBRANE, GLYCOPROTEIN
54	1kb5	A	20	134	1.7e-40			62.17	KB5-C20 T-CELL ANTIGEN RECEPTOR; CHAIN: A, B; ANTIBODY DESIRE-1; CHAIN: L, H;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) TCR VAPLHA VBETA DOMAIN; T-CELL RECEPTOR, STRAND SWITCH, FAB, ANTICLONOTYPIC, 2 (IMMUNOGLOBULIN/RECEPTOR)
54	1kb5	A	20	147	1.7e-40	0.24	0.11		KB5-C20 T-CELL ANTIGEN RECEPTOR; CHAIN: A, B; ANTIBODY DESIRE-1; CHAIN: L, H;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) TCR VAPLHA VBETA DOMAIN; T-CELL RECEPTOR, STRAND SWITCH, FAB, ANTICLONOTYPIC, 2 (IMMUNOGLOBULIN/RECEPTOR)
54	1qrm	D	21	204	1.7e-36			71.85	MHC CLASS I HLA-A; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE P6A;	IMMUNE SYSTEM HUMAN TCR/PEPTIDE/MHC COMPLEX, HLA-A2, HTLV-1, TAX, TCR, T 2 CELL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									CHAIN: C; HMAN T-CELL RECEPTOR; CHAIN: D; HLA-A 0201; CHAIN: E;	RECEPTOR, IMMUNE SYSTEM
54	1qm	D	22	196	1.7e-36	0.38	1.00		MHC CLASS I HLA-A; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE P6A; CHAIN: C; HMAN T-CELL RECEPTOR; CHAIN: D; HLA-A 0201; CHAIN: E;	IMMUNE SYSTEM HUMAN TCR/PEPTIDE/MHC COMPLEX, HLA-A2, HIL V-1, TAX, TCR, T 2 CELL RECEPTOR, IMMUNE SYSTEM
54	lrvf	L	19	131	1.7e-31			51.58	HUMAN RHINOVIRUS 14 COAT PROTEIN; CHAIN: 1, 2, 3, 4; FAB 17-1A; CHAIN: L, H	COMPLEX (COAT PROTEIN/IMMUNOGLOBULIN) POLYPROTEIN, COAT PROTEIN, CORE PROTEIN, RNA-DIRECTED RNA 2 POLYMERASE, HYDROLASE, THIOL PROTEASE, MYRISTYLATION, 3 COMPLEX (COAT PROTEIN/IMMUNOGLOBULIN)
54	litr	A	21	204	1.3e-19			55.23	ALPHA, BETA T-CELL RECEPTOR CHAIN: A, B;	RECEPTOR TCR; T-CELL, RECEPTOR, TRANSMEMBRANE, GLYCOPROTEIN, SIGNAL
54	2fb4	H	19	163	6.8e-36	-0.12	0.55		IMMUNOGLOBULIN IMMUNOGLOBULIN FAB 2FB4 4	
54	8fab	B	19	203	3.4e-32			50.44	IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	
56	1azw	A	107	205	5.1e-05	0.10	0.30		PROLINE IMINOPEPTIDASE; CHAIN: A, B;	AMINOPEPTIDASE AMINOPEPTIDASE, PROLINE IMINOPEPTIDASE, SERINE PROTEASE, 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
56	1brt		80	215	0.00051	0.17	0.15		BROMOPEROXIDASE A2; CHAIN: NULL;	XANTHOMONAS CAMPESTRIS HALOPEROXIDASE HALOPEROXIDASE A2; CHLOROPEROXIDASE A2; HALOPEROXIDASE, OXIDOREDUCTASE, PEROXIDASE, ALPHA/BETA 2 HYDROLASE FOLD, MUTANT M99T
56	1c4x	A	105	216	5.1e-07	0.21	0.04		2-HYDROXY-6-OXO-6- PHENYLHEXA-2,4- DIENOATE CHAIN: A;	HYDROLASE BPHD; HYDROLASE, PCB DEGRADATION
56	1cle	A	1	515	8.5e-57			63.65	CHOLESTEROL ESTERASE; 1CLE 4 CHAIN: A, B; 1CLE 5	LIPASE ESTERASE, SUBSTRATE/PRODUCT- BOUND 1CLE 9
56	1evq	A	411	509	8.5e-13	-0.18	0.36		SERINE HYDROLASE; CHAIN: A;	HYDROLASE ALPHA/BETA
56	1evq	A	73	345	1.7e-39	0.14	1.00		SERINE HYDROLASE; CHAIN: A;	HYDROLASE ALPHA/BETA HYDROLASE FOLD
56	1f6w	A	415	513	5.1e-15	0.05	-0.14		BILE SALT ACTIVATED LIPASE; CHAIN: A;	HYDROLASE BILE SALT ACTIVATED LIPASE, ESTERASE, CATALYTIC DOMAIN
56	1jfr	A	93	299	5.1e-06	-0.35	0.21		LIPASE; CHAIN: A, B;	SERINE HYDROLASE SERINE HYDROLASE, LIPASE
56	1jkm	A	65	321	1.4e-25	-0.16	0.55		BREFELDIN A ESTERASE; CHAIN: A, B;	SERINE HYDROLASE SERINE HYDROLASE, DEGRADATION OF BREFELDIN A, ALPHA/BETA
56	1jkm	A	8	403	1.4e-25			87.04	BREFELDIN A ESTERASE; CHAIN: A, B;	2 HYDROLASE FAMILY SERINE HYDROLASE SERINE HYDROLASE, DEGRADATION OF BREFELDIN A, ALPHA/BETA
56	1lpb	B	107	246	0.00026	0.18	0.07		HYDROLASE(CARBOXYLI C ESTERASE) LIPASE	2 HYDROLASE FAMILY

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
56	1lpp		1	486	5.1e-54			64.96	(E.C.3.1.1.3) COMPLEXED WITH COLIPASE AND INHIBITED ILPB 3 BY UNDECANE PHOSPHONATE METHYL ESTER (TWO CONFORMATIONS) ILPB 4 HYDROLASE LIPASE (E.C.3.1.1.3) (TRIACYLGLYCEROL LIPASE) COMPLEXED WITH ILPP 3 HEXADECANESULFONAT E ILPP 4 ILPP 71	
56	1maa	A	1	515	6.8e-87			62.20	ACETYLCHOLINESTERASE; E; CHAIN: A, B, C, D;	HYDROLASE MACHE; HYDROLASE, SERINE ESTERASE, ACETYLCHOLINESTERASE, TETRAMER, 2 HYDROLASE FOLD, GLYCOSYLATED PROTEIN
56	1qe3	A	23	511	1.7e-67			66.32	PARA-NITROBENZYL ESTERASE; CHAIN: A;	HYDROLASE PNB ESTERASE; ALPHA-BETA HYDROLASE DIRECTED EVOLUTION
56	1qtr	A	65	205	1.2e-05	-0.31	0.40		PROLYL AMINOPEPTIDASE; CHAIN: A;	HYDROLASE ALPHA BETA HYDROLASE FOLD, PROLINE, PROLYL AMINOPEPTIDASE, 2 SERRATIA, IMINOPEPTIDASE
56	1thg		19	197	7.8e-22	-0.37	0.24		HYDROLASE(CARBOXYLI C ESTERASE) LIPASE (E.C.3.1.1.3) TRIACYLGLYCEROL HYDROLASE ITHG 3	
56	1thg		1	497	1.7e-55			75.37	HYDROLASE(CARBOXYLI C ESTERASE) LIPASE (E.C.3.1.1.3)	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
56	2bce		16	514	5.1e-76			58.21	TRIACYLGLYCEROL HYDROLASE 1THG 3 CHOLESTEROL ESTERASE; CHAIN: NULL;	HYDROLASE BILE SALT ACTIVATED LIPASE, BILE SALT STIMULATED HYDROLASE, SERINE ESTERASE, LIPASE
58	1xbr	A	99	306	5.1e-84			161.28	T PROTEIN; CHAIN: A, B; DNA; CHAIN: C, D;	COMPLEX (TRANSCRIPTION FACTOR/DNA) COMPLEX (TRANSCRIPTION FACTOR/DNA), TRANSCRIPTION FACTOR, 2 DNA-BINDING PROTEIN
65	1oqm	A	1	99	6.8e-22	0.35	0.80		RIBOSOMAL PROTEIN S6; CHAIN: A, B;	RIBOSOMAL PROTEIN ALZHEIMER DISEASE, RIBOSOMAL PROTEIN S6, OLIGOMERIZATION
65	1qjh	A	1	95	1.7e-19	0.52	0.92		RIBOSOMAL PROTEIN S6; CHAIN: A;	RIBOSOMAL PROTEIN ALZHEIMER DISEASE, RIBOSOMAL PROTEIN S6, OLIGOMERIZATION
66	1pdr		153	255	8.4e-17	0.48	0.83		HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	SIGNAL TRANSDUCTION HDLG, DHR3 DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, REPEAT
66	1qau	A	157	270	1.2e-17	0.47	0.70		NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: A;	OXIDOREDUCTASE BETA-FINGER
66	1qav	A	153	248	1.2e-17	0.62	0.55		ALPHA-1 SYNTROPHIN (RESIDUES 77-171); CHAIN: A; NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: B;	MEMBRANE PROTEIN/OXIDOREDUCTASE BETA-FINGER, HETERODIMER
66	1qlc	A	157	249	3.6e-18	0.49	0.55		POSTSYNAPTIC DENSITY	PEPTIDE RECOGNITION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									PROTEIN 95; CHAIN: A;	PSD-95; PDZ DOMAIN, NEURONAL NITRIC OXIDE SYNTHASE, NMDA RECEPTOR 2 BINDING
69	1aut	L	40	133	8.4e-09	0.15	-0.17		ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE, PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
69	1c2a	A	35	151	1.2e-22	0.10	-0.07		BOWMAN-BIRK TRYPSIN INHIBITOR; CHAIN: A	HYDROLASE INHIBITOR ALL-BETA STRUCTURE, HYDROLASE INHIBITOR
69	1c2a	A	4	126	4.8e-20	0.84	-0.12		BOWMAN-BIRK TRYPSIN INHIBITOR; CHAIN: A	HYDROLASE INHIBITOR ALL-BETA STRUCTURE, HYDROLASE INHIBITOR
69	1ehd	A	10	110	3.6e-18	0.80	0.30		AGGLUTININ ISOLECTIN VI; CHAIN: A	PLANT PROTEIN TWO HOMOLOGOUS HEVEIN-LIKE DOMAINS
69	1ehd	A	1	30	4.8e-07	2.11	0.48		AGGLUTININ ISOLECTIN VI; CHAIN: A	PLANT PROTEIN TWO HOMOLOGOUS HEVEIN-LIKE DOMAINS
69	1ehd	A	3	87	1.1e-18	1.23	-0.18		AGGLUTININ ISOLECTIN VI; CHAIN: A	PLANT PROTEIN TWO HOMOLOGOUS HEVEIN-LIKE DOMAINS
69	1ehd	A	57	134	4.8e-17	0.75	0.35		AGGLUTININ ISOLECTIN VI; CHAIN: A	PLANT PROTEIN TWO HOMOLOGOUS HEVEIN-LIKE DOMAINS
69	1ehd	A	70	146	1.2e-12	0.22	0.11		AGGLUTININ ISOLECTIN VI; CHAIN: A	PLANT PROTEIN TWO HOMOLOGOUS HEVEIN-LIKE DOMAINS
69	1eis	A	4	87	6e-20	0.98	-0.17		AGGLUTININ ISOLECTIN VI/AGGLUTININ ISOLECTIN V; CHAIN: A;	SUGAR BINDING PROTEIN UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPEROXIDANTIGEN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast 6e-14	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
69	1ext	A	2	155				54.40	TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN, BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
69	1igr	A	4	149	2.4e-22	0.60	-0.19		INSULIN-LIKE GROWTH FACTOR RECEPTOR 1; CHAIN: A;	HORMONE RECEPTOR, HORMONE RECEPTOR, INSULIN RECEPTOR FAMILY
69	1klo		3	143	2.4e-23	0.84	-0.03		LAMININ; CHAIN: NULL;	GLYCOPROTEIN
69	1klo		4	145	2.4e-23			75.71	LAMININ; CHAIN: NULL;	GLYCOPROTEIN
69	1qub	A	4	151	4.8e-24	0.54	-0.20		HUMAN BETA2-GLYCOPROTEIN 1; CHAIN: A;	MEMBRANE ADHESION SHORT CONSENSUS REPEAT, SUSHL, COMPLEMENT CONTROL PROTEIN, 2 N-GLYCOSYLATION, MULTI-DOMAIN, MEMBRANE ADHESION
69	1skz		13	112	6e-17	0.40	-0.17		ANTISTASIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR; ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2 SERINE PROTEASE INHIBITOR, THROMBOSIS
69	1skz		23	124	9.6e-19			58.52	ANTISTASIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR; ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2 SERINE PROTEASE INHIBITOR, THROMBOSIS
69	1skz		30	149	9.6e-19	0.02	-0.05		ANTISTASIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR; ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
69	1vmo	A	2	147	4.8e-28	0.46	-0.18		MEMBRANE PROTEIN VITELLINE MEMBRANE OUTER LAYER PROTEIN I 1VMO 3	SERINE PROTEASE INHIBITOR, THROMBOSIS
69	9wga	A	2	151	3.6e-25	0.56	-0.11		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
69	9wga	A	2	151	3.6e-25			97.03	LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
72	1ee4	A	67	247	0.00024	0.08	0.55		KARYOPHERIN ALPHA; CHAIN: A, B; MYC PROTO- ONCOGENE PROTEIN; CHAIN: C, D, E, F;	TRANSPORT PROTEIN SERINE-RICH RNA POLYMERASE I SUPPRESSOR PROTEIN; ARM REPEAT
80	1fch	A	371	574	0.00024	-0.23	0.96		PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR; CHAIN: A, B; PTS1-CONTAINING PEPTIDE; CHAIN: C, D;	SIGNALING PROTEIN PEROXISOME RECEPTOR 1, PTS1-BP, PEROXIN-5, PTS1 PROTEIN-PEPTIDE COMPLEX, TETRATRICOPEPTIDE REPEAT, TPR, 2 HELICAL REPEAT
83	1cs6	A	34	204	6e-15	0.22	0.68		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
83	1cvs	C	34	202	1.2e-13	-0.09	0.90		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN- LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR

Table 5

SEQ ID NO.	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
83	1epf	A	34	202	1.2e-15	0.34	0.92		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	RECEPTOR CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
83	1ev2	E	28	202	1.1e-15	0.12	0.31		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGF2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
83	1ev2	G	34	203	3.6e-13	0.12	0.65		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGF2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
83	1f2q	A	34	193	3.6e-13	0.04	0.34		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RI-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
83	1f6a	A	34	202	6e-14	-0.14	0.27		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IG EPSILON CHAIN C REGION; CHAIN: B, D;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
83	1fcg	A	34	203	4.8e-15	0.04	0.29		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32, FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
83	2fcb	A	34	202	3.6e-15	0.33	0.71		FC GAMMA RUBB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
83	2ncm		104	202	2.4e-14	0.49	0.94		NEURAL CELL ADHESION MOLECULE; CHAIN: NULL;	IMMUNE SYSTEM CELL ADHESION NCAM DOMAIN 1; CELL ADHESION, GLYCOPROTEIN, HEPARIN-BINDING, GPI-ANCHOR, 2 NEURAL ADHESION MOLECULE, IMMUNOGLOBULIN FOLD, SIGNAL
85	1ee4	A	11	225	4.8e-09	0.35	0.68		KARYOPHERIN ALPHA; CHAIN: A, B; MYC PROTO-ONCOGENE PROTEIN; CHAIN: C, D, E, F;	TRANSPORT PROTEIN SERINE-RICH RNA POLYMERASE I SUPPRESSOR PROTEIN; ARM REPEAT
85	1ee4	A	26	311	2.4e-05	0.16	0.80		KARYOPHERIN ALPHA; CHAIN: A, B; MYC PROTO-ONCOGENE PROTEIN; CHAIN: C, D, E, F;	TRANSPORT PROTEIN SERINE-RICH RNA POLYMERASE I SUPPRESSOR PROTEIN; ARM REPEAT
85	1ez3	A	737	867	3.6e-15	0.01	-0.14		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
85	3bet		108	532	2.4e-09	0.20	0.18		BETA-CATENIN; CHAIN: NULL;	ARMADILLO REPEAT ARMADILLO REPEAT, BETA-CATENIN, CYTOSKELETON
85	3bet		7	440	1.2e-09	0.15	0.93		BETA-CATENIN; CHAIN: NULL;	ARMADILLO REPEAT ARMADILLO REPEAT, BETA-CATENIN, CYTOSKELETON
86	1alh	A	156	232	6.8e-28	-0.27	0.59		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									BINDING SITE; CHAIN: B, C;	PROTEIN
86	1alh	A	180	260	3.4e-27	-0.41	1.00		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
86	1alh	A	291	365	8.5e-27	0.10	-0.06		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
86	1alh	A	313	391	6.8e-28	0.02	0.53		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
86	1alh	A	651	733	8.5e-29	-0.37	0.00		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
86	1alh	A	707	785	3.4e-25	-0.21	0.19		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
86	1alh	A	739	814	1.2e-23	-0.44	0.21		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
86	1alh	A	789	870	1e-29	0.11	0.63		QGR ZINC FINGER	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN

Table 5

SEQ ID NO;	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
86	1alh	A	818	899	1.4e-28	-0.05	0.48		PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
86	1alh	A	846	928	7.2e-14	-0.17	0.43		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
86	1alh	A	875	946	5.1e-27	-0.09	0.31		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
86	1ard		313	341	6.8e-06	-0.69	0.00		TRANSCRIPTION REGULATION YEAST TRANSCRIPTION FACTOR ADRI (RESIDUES 102 - 130) IARD 3 (AMINO TERMINAL ZINC FINGER DOMAIN) (NMR, 10 STRUCTURES) IARD 4 (ADRI) IARD 5	
86	1bbo		848	897	3.4e-10	-0.31	0.19		DNA-BINDING PROTEIN HUMAN ENHANCER- BINDING PROTEIN MBP-1 MUTANT WITH CYS 11 IBBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) IBBO 4	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
86	1mey	C	155	232	6.8e-46	-0.13	0.81		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
86	1mey	C	179	260	8.5e-46	0.01	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
86	1mey	C	207	287	1.7e-43	-0.16	0.95		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
86	1mey	C	263	337	3.4e-43	-0.70	0.03		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
86	1mey	C	312	391	5.1e-46	-0.22	0.70		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
86	1mey	C	650	733	1e-46	-0.38	0.35		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
86	1mey	C	706	785	8.5e-41	-0.31	0.36		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
86	1mey	C	766	842	1.5e-42	-0.11	0.28		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
86	1mey	C	788	870	6.8e-49	0.06	0.86		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
86	1mey	C	817	899	3.4e-48	-0.02	0.70		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
86	1mey	G	123	152	1.7e-10	0.18	-0.19		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
86	1mey	G	205	232	5.1e-13	-0.25	0.34		DNA; CHAIN: A, B, D, E;	COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
86	1mey	G	843	870	6.8e-13	0.27	0.17		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2-CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
86	1mey	G	871	899	6.8e-11	0.22	0.65		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
86	1mey	G	947	974	3.4e-12	-0.32	0.10		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
86	1sp2		846	873	1e-07	-0.38	0.03		SP1F2; CHAIN: NULL;	ZINC FINGER TRANSCRIPTION FACTOR SP1; ZINC FINGER, TRANSCRIPTION ACTIVATION, SP1
86	1tf3	A	180	260	3.4e-21	-0.07	0.88		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
86	1tf3	A	707	785	5.1e-18	-0.32	0.13		TRANSCRIPTION FACTOR IIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
86	1tf6	A	707	858	3.4e-33	-0.37	0.47		TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE II, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
86	1tf6	A	818	983	1e-35	-0.21	0.18		TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE II, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
86	1ubd	C	135	232	5.1e-32	-0.11	0.59		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
86	1ubd	C	270	391	5.1e-29	-0.51	0.25		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
86	1ubd	C	630	733	5.1e-35	-0.49	0.07		ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
86	1ubd	C	658	763	3.4e-33	-0.26	0.45		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
86	1ubd	C	711	814	1.7e-31	-0.41	0.83		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
86	1ubd	C	796	899	1e-31	-0.35	0.18		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
86	1ubd	C	853	974	6.8e-32	-0.41	0.13		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
86	2adr		237	289	5.1e-13	-0.62	0.80		ADRI; CHAIN: NULL;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
86	2adr		707	765	1.2e-12	-0.10	0.22		ADRI; CHAIN: NULL;	TRANSCRIPTION REGULATION REGULATION, ADRI, ZINC FINGER, NMR
86	2adr		846	901	5.1e-15	-0.29	0.10		ADRI; CHAIN: NULL;	TRANSCRIPTION REGULATION REGULATION, ADRI, ZINC FINGER, NMR
86	2drp	A	844	898	1e-08	-0.06	0.07		COMPLEX(TRANSCRIPTION REGULATION/DNA) TRAMTRACK PROTEIN (TWO ZINC-FINGER PEPTIDE) COMPLEXED WITH 2DRP 3 DNA 2DRP 4	
86	2gli	A	135	260	1.5e-32	-0.04	0.89		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
86	2gli	A	180	337	1.2e-21	-0.27	0.07		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	BINDING PROTEIN/DNA) COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
86	2gli	A	184	285	2.4e-23	-0.57	0.33		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
86	2gli	A	207	365	1.2e-16	-0.05	0.11		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
86	2gli	A	270	393	1.4e-31	-0.37	0.12		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
86	2gli	A	630	765	1.7e-32	-0.36	0.29		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
86	2gli	A	739	899	4.8e-21	-0.44	0.45		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
86	2gli	A	772	901	8.5e-31	-0.28	0.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
86	2gli	A	792	923	2.4e-21	-0.39	0.22		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
86	2gli	A	818	931	2.4e-20	0.11	0.70		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
86	2gli	A	825	976	8.5e-31	-0.31	0.27		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
87	1bih	A	23	518	3.6e-21	0.06	0.09		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
87	1bih	A	275	796	2.4e-11	-0.07	0.16		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
87	1cdy		160	376	2.4e-11	-0.04	0.36		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: NULL;	T-CELL SURFACE GLYCOPROTEIN IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2
87	1cdy		30	250	8.4e-09	0.18	0.29		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: NULL;	MHC, LIPOPROTEIN, T-CELL SURFACE GLYCOPROTEIN
87	1cs6	A	22	518	1.1e-16	-0.11	0.36		AXONIN-1; CHAIN: A;	T-CELL SURFACE GLYCOPROTEIN IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2
87	1cs6	A	284	796	2.4e-09	-0.18	0.29		AXONIN-1; CHAIN: A;	MHC, LIPOPROTEIN, T-CELL SURFACE GLYCOPROTEIN
87	1cvs	C	104	250	2.4e-11	-0.33	0.22		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B;	CELL ADHESION NEURAL CELL ADHESION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
87	levs	C	145	376	2.4e-12	0.06	0.29		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
87	lepf	A	153	376	1.2e-10	0.11	0.25		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
87	lepf	A	24	269	1.2e-07	-0.03	0.49		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
87	lepf	A	284	541	3.6e-10	0.05	0.07		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
87	lepf	A	471	658	8.4e-05	-0.18	0.18		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
87	lev2	E	154	376	3.6e-11	-0.25	0.15		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
87	lev2	G	104	250	9.6e-09	-0.29	0.17		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
87	1ev2	G	153	376	4.8e-10	-0.10	0.24		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	DOMAINS, B-TREFOIL FOLD GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
87	1evt	C	104	250	7.2e-08	-0.54	0.05		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
87	1evt	C	143	376	8.4e-09	-0.08	0.00		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
87	1f2q	A	104	250	1.2e-08	-0.24	0.06		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RI-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
87	1f2q	A	154	376	2.4e-05	-0.08	0.31		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RI-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
87	1f6a	A	284	541	0.00024	-0.18	0.25		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IG EPSILON CHAIN C REGION; CHAIN:	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									B, D;	RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
87	1fcg	A	282	541	0.0024	-0.06	0.06		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
87	1fyt	D	284	383	0.0048	0.27	0.88		HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DR CHAIN: A; HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DR-1 CHAIN: B; HEMAGGLUTININ HA1 PEPTIDE CHAIN; CHAIN: C; T-CELL RECEPTOR ALPHA CHAIN; CHAIN: D; T-CELL RECEPTOR BETA CHAIN; CHAIN: E;	IMMUNE SYSTEM HLA-DRI, DRA; HLA-DRI, DRB1 0101; TCR HA1.7 ALPHA CHAIN; TCR HA1.7 BETA CHAIN; PROTEIN-PROTEIN COMPLEX, IMMUNOGLOBULIN FOLD
87	1neu		23	120	8.4e-06	-0.24	0.01		MYELIN P0 PROTEIN; CHAIN: NULL;	STRUCTURAL PROTEIN MYELIN, STRUCTURAL PROTEIN, GLYCOPROTEIN, TRANSMEMBRANE, PHOSPHORYLATION, IMMUNOGLOBULIN FOLD, SIGNAL, MYELIN 2 MEMBRANE ADHESION MOLECULE
87	1neu		417	517	6e-05	0.43	0.55		MYELIN P0 PROTEIN; CHAIN: NULL;	STRUCTURAL PROTEIN MYELIN, STRUCTURAL PROTEIN, GLYCOPROTEIN, TRANSMEMBRANE, PHOSPHORYLATION, IMMUNOGLOBULIN FOLD, SIGNAL, MYELIN 2 MEMBRANE ADHESION MOLECULE
87	1qfo	A	283	377	0.00036	0.06	0.46		SIALOADHESIN; CHAIN:	IMMUNE SYSTEM

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
87	1wio	A	160	486	6e-10	0.06	0.52		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	IMMUNOGLOBULIN SUPERFAMILY, CARBOHYDRATE BINDING GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC LIPOPROTEIN, POLYMORPHISM
87	1wio	A	30	380	6e-14	-0.13	0.10		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC LIPOPROTEIN, POLYMORPHISM
87	2feb	A	104	250	2.4e-10	-0.11	0.11		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
87	2feb	A	284	518	0.0012	-0.24	0.04		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
87	2ncm		22	119	0.0024	-0.06	0.28		NEURAL CELL ADHESION MOLECULE; CHAIN: NULL;	CELL ADHESION NCAM DOMAIN 1; CELL ADHESION, GLYCOPROTEIN, HEPARIN-BINDING, GPI-ANCHOR, 2 NEURAL ADHESION MOLECULE, IMMUNOGLOBULIN FOLD, SIGNAL
87	2ncm		284	401	6e-06	0.28	0.22		NEURAL CELL ADHESION MOLECULE; CHAIN: NULL;	CELL ADHESION NCAM DOMAIN 1; CELL ADHESION, GLYCOPROTEIN, HEPARIN-BINDING, GPI-ANCHOR, 2 NEURAL ADHESION MOLECULE, IMMUNOGLOBULIN FOLD, SIGNAL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
90	1bkf		56	163	8.4e-44			95.53	FK506 BINDING PROTEIN; CHAIN: NULL;	ISOMERASE FKBP; ISOMERASE, ROTAMASE
90	1bkf		60	163	8.4e-44	0.99	1.00		FK506 BINDING PROTEIN; CHAIN: NULL;	ISOMERASE FKBP; ISOMERASE, ROTAMASE
90	1c9h	A	57	163	6e-44	0.82	1.00		FKBP12.6; CHAIN: A;	IMMUNE SYSTEM CALCINEURIN; FKBP12; RAPAMYCIN, COMPLEX, RYANODINE RECEPTOR
90	1rot		50	164	3.6e-41			93.58	FKBP59-I; CHAIN: NULL;	ROTAMASE (ISOMERASE) FKBP52 OR HSP56; ROTAMASE (ISOMERASE), DOMAIN 1 (N-TERM) OF A 59 KDA, 2 FK506-BINDING PROTEIN, PEPTIDYL PROLYL CIS-TRANS ISOMERASE
90	1rot		56	163	3.6e-41	0.82	1.00		FKBP59-I; CHAIN: NULL;	ROTAMASE (ISOMERASE) FKBP52 OR HSP56; ROTAMASE (ISOMERASE), DOMAIN 1 (N-TERM) OF A 59 KDA, 2 FK506-BINDING PROTEIN, PEPTIDYL PROLYL CIS-TRANS ISOMERASE
94	1ern	A	665	737	2.4e-05	-0.20	0.45		ERYTHROPOIETIN RECEPTOR; CHAIN: A, B;	CYTOKINE EBP; ERYTHROPOIETIN RECEPTOR, SIGNAL TRANSDUCTION, CYTOKINE 2 RECEPTOR CLASS 1
94	166f	B	666	734	2.4e-05	-0.47	0.16		PLACENTAL LACTOGEN; CHAIN: A; PROLACTIN RECEPTOR; CHAIN: B, C;	HORMONE/GROWTH FACTOR/HORMONE RECEPTOR 4-HELICAL BUNDLE, ALPHA HELICAL BUNDLE, TERNARY COMPLEX, FN 2 III DOMAINS, BETA SHEET

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
94	1ten		659	732	0.0012	-0.52	0.05		CELL ADHESION PROTEIN TENASCIN (THIRD FIBRONECTIN TYPE III REPEAT) 1TEN 3	DOMAINS, CYTOKINE-RECEPTOR COMPLEX
95	1qbx	A	393	513	0.0024	-0.07	0.07		CHLORAMPHENICOL PHOSPHOTRANSFERASE; CHAIN: A;	TRANSFERASE CPT; KINASE, ANTIBIOTIC RESISTANCE, PHOSPHORYLATION, 2 MONONUCLEOTIDE BINDING FOLD
96	1qde	A	1360	1635	4.8e-11	0.02	-0.18		VESICULAR TRANSPORT PROTEIN SEC17; CHAIN: A;	PROTEIN TRANSPORT HELIX-TURN-HELIX TPR-LIKE REPEAT, PROTEIN TRANSPORT
97	1bih	A	18	387	1.2e-29			91.03	HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
97	1bih	A	30	297	1.2e-29	0.27	0.39		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
97	1cdy		22	105	0.00096	-0.18	0.03		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: NULL;	T-CELL SURFACE GLYCOPROTEIN IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 25 MHC, LIPOPROTEIN, T-CELL SURFACE GLYCOPROTEIN
97	1cs6	A	29	312	2.4e-30	-0.12	0.76		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
97	1cs6	A	6	202	3.6e-18	0.24	0.01		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
97	1cvs	C	109	289	1.2e-21	-0.00	0.06		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
97	1cvs	C	30	190	3.6e-14	-0.15	0.71		FACTOR RECEPTOR 1; CHAIN: C, D;	LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
97	1e4k	A	10	147	0.0011	-0.00	0.12		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
97	1e4k	A	113	302	6e-07	0.39	0.29		LOW AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR CHAIN: C; FC FRAGMENT OF HUMAN IGG1; CHAIN: A, B;	COMPLEX CD16; IGG1-FC COMPLEX, FC FRAGMENT, IGG, FC, RECEPTOR, CD16, GAMMA
97	1epf	A	118	285	6e-22	0.12	0.09		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
97	1epf	A	15	105	2.4e-06	0.57	0.34		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
97	1epf	A	215	312	3.6e-08	0.10	-0.01		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
97	1epf	A	30	205	3.6e-13	0.20	0.49		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
97	1ev2	E	107	289	7.2e-22	0.06	0.55		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; CHAIN: A, B, C;

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
97	1ev2	E	15	98	0.0006	0.01	0.63		D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
97	1ev2	E	29	190	3.6e-16	-0.05	0.16		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
97	1ev2	G	107	297	1.2e-24	0.12	0.46		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
97	1ev2	G	29	204	3.6e-18	-0.00	0.09		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
97	1ev1	C	107	289	2.4e-23	-0.06	0.12		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
97	1f2q	A	115	297	1.2e-21	0.31	0.71		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	DOMAINS, B-TREFOIL FOLD IMMUNE SYSTEM FC-EPSILON RI-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
97	1f2q	A	16	109	1.2e-08	0.36	0.36		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RI-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
97	1f2q	A	30	204	6e-23	0.30	0.30		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RI-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
97	1f6a	A	108	297	1.2e-23	0.35	0.72		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IG EPSILON CHAIN C REGION; CHAIN: B, D;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
97	1f6a	A	16	204	4.8e-24	0.25	0.88		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IG EPSILON CHAIN C REGION; CHAIN: B, D;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
97	1f6g	A	115	292	2.4e-25	0.34	0.22		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD3241 FC RECEPTOR, IMMUNOGLOBULIN,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
97	1feg	A	16	203	1.1e-23	0.06	0.40		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	LEUKOCYTE, CD32 IMMUNE SYSTEM. MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
97	1feg	A	208	297	8.4e-09	0.43	0.69		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM. MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
97	1hmf		25	89	6e-05	0.17	0.12		T LYMPHOCYTE ADHESION GLYCOPROTEIN CD2 (HUMAN) IHNF 3	
97	1hmf		30	187	3.6e-12	-0.03	0.10		T LYMPHOCYTE ADHESION GLYCOPROTEIN CD2 (HUMAN) IHNF 3	
97	1nkr		120	276	4.8e-15	0.33	0.13		P58-CL42 KIR; CHAIN: NULL;	INHIBITORY RECEPTOR KILLER CELL INHIBITORY RECEPTOR; INHIBITORY RECEPTOR, NATURAL KILLER CELLS, IMMUNOLOGICAL 2 RECEPTORS, IMMUNOGLOBULIN FOLD
97	1nkr		20	201	6e-12	0.20	0.13		P58-CL42 KIR; CHAIN: NULL;	INHIBITORY RECEPTOR KILLER CELL INHIBITORY RECEPTOR; INHIBITORY RECEPTOR, NATURAL KILLER CELLS, IMMUNOLOGICAL 2 RECEPTORS, IMMUNOGLOBULIN FOLD
97	1wfo	A	39	376	9.6e-14			73.40	T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
97	2feb	A	114	297	1.2e-28	0.23	0.64		FC GAMMA RIB; CHAIN: A;	MHC LIPOPROTEIN, POLYMORPHISM
97	2feb	A	16	204	2.4e-22	0.23	0.49		FC GAMMA RIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
98	1bih	A	11	380	1.2e-29			89.98	HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
98	1bih	A	23	290	1.2e-29	0.27	0.39		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
98	1cdy		15	98	0.00096	-0.18	0.03		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: NULL;	T-CELL SURFACE GLYCOPROTEIN IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC, LIPOPROTEIN, T-CELL SURFACE GLYCOPROTEIN
98	1cs6	A	22	305	2.4e-30	-0.12	0.76		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
98	1evs	C	102	282	1.2e-21	-0.00	0.06		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
98	1e4k	A	106	295	6e-07	0.39	0.29		LOW AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR CHAIN: C; FC FRAGMENT OF HUMAN IGG1; CHAIN: A, B;	COMPLEX CD16; IGG1-FC COMPLEX, FC FRAGMENT, IGG, FC, RECEPTOR, CD16, GAMMA
98	1e4k	A	24	140	0.0036	-0.31	0.46		LOW AFFINITY	COMPLEX CD16; IGG1-FC

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									IMMUNOGLOBULIN GAMMA FC RECEPTOR CHAIN: C; FC FRAGMENT OF HUMAN IGG1; CHAIN: A, B;	COMPLEX, FC FRAGMENT, IGG, FC, RECEPTOR, CD16, GAMMA
98	1epf	A	111	278	6e-22	0.12	0.09		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
98	1epf	A	17	98	4.8e-06	0.35	0.46		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
98	1epf	A	208	305	3.6e-08	0.10	-0.01		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
98	1ev2	E	100	282	7.2e-22	0.06	0.55		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGF2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
98	1ev2	E	18	91	0.0012	-0.06	0.48		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGF2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
98	1ev2	E	22	183	3.6e-16	-0.05	0.16		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGF2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
98	1ev2	G	100	290	1.2e-24	0.12	0.46		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGF2; IMMUNOGLOBULIN (IG) LIKE DOMAINS

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
98	1ev2	G	22	197	3.6e-18	-0.00	0.09		CHAIN: E, F, G, H; FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGF2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
98	1evt	C	100	282	2.4e-23	-0.06	0.12		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1; FGF1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
98	1f2q	A	108	290	1.2e-21	0.31	0.71		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RI-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
98	1f2q	A	23	197	6e-23	0.30	0.30		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RI-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
98	1f6a	A	101	290	1.2e-23	0.35	0.72		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IG EPSILON CHAIN C REGION; CHAIN: B, D;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
98	1f6a	A	14	197	7.2e-24	0.33	0.13		HIGH AFFINITY IMMUNOGLOBULIN	IMMUNE SYSTEM HIGH AFFINITY IGE-FC

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
98	1fcg	A	108	285	2.4e-25	0.34	0.22		EPSILON RECEPTOR CHAIN: A; IG EPSILON CHAIN C REGION; CHAIN: B, D;	RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR; IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
98	1fcg	A	12	99	1.2e-06	0.49	0.10		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
98	1fcg	A	17	196	2.4e-23	0.05	0.10		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
98	1fcg	A	201	290	8.4e-09	0.43	0.69		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
98	1hnf		18	82	6e-05	0.17	0.12		T LYMPHOCYTE ADHESION GLYCOPROTEIN CD2 (HUMAN) IHN3	
98	1hnf		23	180	3.6e-12	-0.03	0.10		T LYMPHOCYTE ADHESION GLYCOPROTEIN CD2 (HUMAN) IHN3	
98	1nkr		113	269	4.8e-15	0.33	0.13		P58-CL42 KIR; CHAIN: NULL;	INHIBITORY RECEPTOR KILLER CELL INHIBITORY RECEPTOR; INHIBITORY RECEPTOR, NATURAL KILLER CELLS,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
98	1wio	A	32	369	9.6e-14			72.45	T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	IMMUNOLOGICAL 2 RECEPTORS, IMMUNOGLOBULIN FOLD
98	2feb	A	107	290	1.2e-28	0.23	0.64		FC GAMMA RIIB; CHAIN: A;	GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC LIPOPROTEIN, POLYMORPHISM
98	2feb	A	17	197	4.8e-22	0.34	0.45		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
537	1fjs	L	1068	1104	0.00072	0.23	0.00		COAGULATION FACTOR XA; CHAIN: A; COAGULATION FACTOR XA; CHAIN: L;	BLOOD CLOTTING PROTEIN INHIBITOR COMPLEX, COAGULATION COFACTOR, PROTEASE
102	1alh	A	182	260	4.8e-27	-0.09	0.74		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
102	1alh	A	209	288	1.1e-37	0.27	1.00		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
102	1alh	A	235	316	7.2e-42	0.64	1.00		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
102	1alh	A	263	344	6e-45	0.38	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
102	1alh	A	291	372	3.6e-46	0.10	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
102	1alh	A	319	400	7.2e-46	0.34	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
102	1alh	A	347	428	1.2e-44	0.02	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
102	1alh	A	375	456	2.4e-45	0.46	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
102	1alh	A	403	484	3.6e-46	-0.11	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
102	1alh	A	431	511	2.4e-46	0.10	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
102	1alh	A	459	532	1.2e-31	-0.17	1.00		OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	FINGER, DNA-BINDING PROTEIN
102	1alh	A	564	638	3.6e-27	0.11	0.39		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
102	1alh	A	592	649	1.2e-21	0.33	0.71		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
102	1ard		209	235	2.4e-10	0.10	0.94		TRANSCRIPTION REGULATION YEAST TRANSCRIPTION FACTOR ADRI (RESIDUES 102 - 130) LARD 3 (AMINO TERMINAL ZINC FINGER DOMAIN) (NMR, 10 STRUCTURES) LARD 4 (ADRI) LARD 5	
102	1ard		487	514	8.4e-10	0.83	0.77		TRANSCRIPTION REGULATION YEAST TRANSCRIPTION FACTOR ADRI (RESIDUES 102 - 130) LARD 3 (AMINO TERMINAL ZINC FINGER DOMAIN) (NMR, 10 STRUCTURES) LARD 4 (ADRI) LARD 5	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
102	1bbo		461	516	6e-27	-0.32	0.13		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11 IBBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) IBBO 4	
102	1bbo		566	620	2.4e-19	0.02	-0.01		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11 IBBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) IBBO 4	
102	1bbo		594	648	2.4e-19	-0.36	0.12		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11 IBBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) IBBO 4	
102	1paa		487	513	4.8e-07	0.38	0.99		TRANSCRIPTION REGULATION YEAST TRANSCRIPTION FACTOR ADRI (RESIDUES 130 - 159) IPAA 3 (PAPA - CARBOXY TERMINAL ZINC FINGER DOMAIN) MUTANT WITH IPAA 4 PRO 131 REPLACED BY ALA, PRO 133 REPLACED BY ALA, CYS 140 IPAA 5 REPLACED BY ALA (PI31A.PI33A.C140A) (NMR, 10 STRUCTURES) IPAA 6	
102	1sp1		487	514	6e-09	0.15	0.18		SPIF3; CHAIN: NULL;	ZINC FINGER TRANSCRIPTION FACTOR SPI; ZINC FINGER,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
102	1sp2		291	319	3.6e-11	0.49	0.59		SPIF2; CHAIN: NULL;	TRANSCRIPTION ACTIVATION, SP1 ZINC FINGER TRANSCRIPTION FACTOR SPI; ZINC FINGER, TRANSCRIPTION ACTIVATION, SP1
102	1sp2		431	459	1.2e-12	0.37	0.99		SPIF2; CHAIN: NULL;	ZINC FINGER TRANSCRIPTION FACTOR SPI; ZINC FINGER, TRANSCRIPTION ACTIVATION, SP1
102	1sp2		487	514	3.6e-09	0.36	0.53		SPIF2; CHAIN: NULL;	ZINC FINGER TRANSCRIPTION FACTOR SPI; ZINC FINGER, TRANSCRIPTION ACTIVATION, SP1
102	1tf3	A	182	261	6e-16	-0.05	0.11		TRANSCRIPTION FACTOR III A; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
102	1tf3	A	209	290	6e-27	0.07	0.51		TRANSCRIPTION FACTOR III A; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
102	1tf3	A	263	346	4.8e-33	0.51	0.99		TRANSCRIPTION FACTOR III A; CHAIN: A; 5S RNA	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA;

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									GENE; CHAIN: E, F;	5S GENE; NMR, TFIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
102	1tf3	A	459	541	6e-21	0.05	0.41		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIA; 5S GENE; NMR, TFIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
102	1tf3	A	564	638	2.4e-15	0.20	-0.03		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIA; 5S GENE; NMR, TFIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
102	1tf3	A	592	649	9.6e-15	0.24	0.34		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIA; 5S GENE; NMR, TFIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
102	2gli	A	182	345	1.2e-54	0.03	0.95		ZINC FINGER PROTEIN	COMPLEX (DNA-BINDING REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									GLI1; CHAIN: A; DNA; CHAIN: C, D;	PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
102	2gli	A	234	373	2.4e-70	0.48	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
102	2gli	A	290	457	2.4e-74	0.19	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
102	2gli	A	318	457	2.4e-74			96.84	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
102	2gli	A	346	513	1.2e-73	-0.00	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
102	2gli	A	430	618	6e-51	0.14	0.60		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
103	1alh	A	123	195	3.4e-24	-0.10	0.25		QGSF ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
103	1alh	A	144	222	8.5e-25	0.03	0.03		QGSF ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
103	1alh	A	325	435	1.2e-42	0.03	0.47		OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	FINGER, DNA-BINDING PROTEIN
103	1mey	C	120	195	8.5e-41	0.13	1.00		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
103	1mey	C	143	222	3.4e-43	0.16	0.98		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
103	1mey	C	198	293	1.4e-43	-0.46	0.89		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
103	1mey	C	240	321	1.7e-49	0.40	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
103	1mey	C	268	349	1.2e-50	0.31	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER	COMPLEX (ZINC FINGER/DNA) ZINC FINGER,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
103	1mey	C	268	350	1.7e-51			99.05	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
103	1mey	C	296	377	1.7e-51	0.42	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
103	1mey	C	324	405	5.1e-51	0.36	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
103	1mey	C	352	433	1e-50	0.47	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
103	1mey	C	380	461	6.8e-51	0.50	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
103	1mey	C	408	489	1e-50	0.80	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	(ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
103	1mey	C	436	504	5.1e-43	0.29	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
103	1mey	G	141	168	3.4e-11	-0.02	0.12		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
103	1mey	G	266	293	1.7e-12	0.94	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
103	1tf3	A	123	195	1.2e-14	-0.23	0.01		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) THIIA; 5S GENE; NMR, THIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
103	1tf3	A	144	218	6.8e-17	0.03	0.06		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
103	1tf6	A	199	358	8.5e-35	-0.20	0.86		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
103	1tf6	A	239	403	5.1e-38			100.94	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
103	1tf6	A	241	386	5.1e-38	0.09	0.99		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
103	1tf6	A	353	503	1.7e-38	0.26	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2

Table 5

SEQ ID No:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
103	1ubd	C	123	223	1e-29	0.08	0.03		YY1; CHAIN: C; ADENO- ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	TRANSCRIPTION, INITIATION, ZINC FINGER PROTEIN
103	1ubd	C	218	321	5.1e-32	0.02	1.00		YY1; CHAIN: C; ADENO- ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING- REGULATION/DNA) YING- YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA- PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
103	1ubd	C	248	349	3.4e-35	0.30	1.00		YY1; CHAIN: C; ADENO- ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING- REGULATION/DNA) YING- YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA- PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
103	1ubd	C	270	378	3.4e-35			84.44	YY1; CHAIN: C; ADENO- ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING- REGULATION/DNA) YING- YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA- PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
103	1ubd	C	276	377	8.5e-35	0.21	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
103	1ubd	C	329	433	3.4e-35	0.26	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
103	1ubd	C	360	461	1.5e-34	0.40	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
103	1ubd	C	388	489	3.4e-35	0.16	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
103	1ubd	C	416	504	1.5e-29	0.28	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									DNA; CHAIN: A, B;	INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
103	2adr		120	170	3.4e-13	-0.14	0.00		ADRI; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADRI, ZINC FINGER, NMR
103	2adr		144	197	1.7e-13	0.02	0.17		ADRI; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADRI, ZINC FINGER, NMR
103	2gli	A	143	351	3.6e-52	-0.40	0.23		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
103	2gli	A	205	348	1.7e-33	0.17	0.90		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
103	2gli	A	244	379	1.2e-66	0.51	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
103	2gli	A	268	407	1.2e-70			92.19	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
103	2gli	A	268	433	1.2e-70	0.29	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
103	2gli	A	304	432	3.4e-33	0.50	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
103	2gli	A	332	460	1e-32	0.38	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
103	2gli	A	352	494	1.2e-52	0.32	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
103	2gli	A	360	488	1.7e-33	0.49	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
103	2gli	A	388	498	6.8e-30	-0.03	0.98		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
103	2gli	A	416	504	3.4e-21	-0.01	0.41		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
103	2gli	A	64	221	8.5e-29	-0.72	0.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
104	1c1g	A	216	442	1.2e-10	0.09	-0.20		TROPOMYOSIN; CHAIN: A;	CONTRACTILE PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									B, C, D	TROPOMYOSIN COILED-COIL ALPHA-HELICAL, CONTRACTILE PROTEIN
104	1c1g	A	261	503	3.4e-11	0.09	-0.19		TROPOMYOSIN; CHAIN: A, B, C, D	CONTRACTILE PROTEIN TROPOMYOSIN COILED-COIL ALPHA-HELICAL, CONTRACTILE PROTEIN
105	1c1g	A	260	503	3.4e-19	0.02	-0.20		TROPOMYOSIN; CHAIN: A, B, C, D	CONTRACTILE PROTEIN TROPOMYOSIN COILED-COIL ALPHA-HELICAL, CONTRACTILE PROTEIN
107	1a1t	A	1347	1383	3.6e-07	-0.33	0.17		NUCLEOCAPSID PROTEIN; CHAIN: A; SL3 STEM-LOOP RNA; CHAIN: B;	COMPLEX (NUCLEOCAPSID PROTEIN/RNA) NUCLEOCAPSID PROTEIN, COMPLEX (NUCLEOCAPSID PROTEIN/RNA), 2 STEM-LOOP RNA
107	1aaf		1347	1383	2.4e-07	-0.53	0.12		NUCLEOCAPSID PROTEIN HIV-1 NUCLEOCAPSID PROTEIN (MN ISOLATE) (NMR, 20 STRUCTURES) 1AAF 3	
107	1f5a	A	343	542	6e-05	-0.04	0.53		POLY(A) POLYMERASE; CHAIN: A;	TRANSFERASE MRNA PROCESSING, TRANSFERASE, TRANSCRIPTION, RNA-BINDING, 2 PHOSPHORYLATION, NUCLEAR PROTEIN, ALTERNATIVE SPLICING 3 HELICAL TURN MOTIF, NUCLEOTIDYL TRANSFERASE CATALYTIC DOMAIN
107	1f5a	A	996	1147	2.4e-07	0.01	0.41		POLY(A) POLYMERASE; CHAIN: A;	TRANSFERASE MRNA PROCESSING, TRANSFERASE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										TRANSCRIPTION, RNA-BINDING, 2 PHOSPHORYLATION, NUCLEAR PROTEIN, ALTERNATIVE SPLICING 3 HELICAL TURN MOTIF, NUCLEOTIDYL TRANSFERASE CATALYTIC DOMAIN
108	1alh	A	257	312	2.4e-27	0.40	0.66		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLICATION OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
108	1ard		259	287	2.4e-10	0.14	0.48		TRANSCRIPTION REGULATION YEAST TRANSCRIPTION FACTOR ADRI (RESIDUES 102 - 130) 1ARD 3 (AMINO TERMINAL ZINC FINGER DOMAIN) (NMR, 10 STRUCTURES) 1ARD 4 (ADRI) 1ARD 5	
108	1bbo		257	287	1.2e-11	0.13	0.18		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11 1BBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) 1BBO 4	
108	1mey	C	262	339	6.8e-47	0.27	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
108	1mey	C	286	367	1.2e-50	0.36	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
108	1mey	C	286	368	1.7e-51			116.59	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
108	1mey	C	314	395	1.7e-51	0.42	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
108	1mey	C	342	424	3.4e-49	0.28	0.89		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
108	1mey	C	370	429	1.7e-34	-0.08	0.16		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
108	1paa		259	285	0.00096	-0.05	0.55		TRANSCRIPTION REGULATION YEAST TRANSCRIPTION FACTOR ADRI (RESIDUES 130 - 159)	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
108	1sp2		259	287	1.1e-09	0.30	0.46		IPAA 3 (PAPA - CARBOXY TERMINAL ZINC FINGER DOMAIN) MUTANT WITH IPAA 4 PRO 131 REPLACED BY ALA, PRO 133 REPLACED BY ALA, CYS 140 IPAA 5 REPLACED BY ALA (P131A,P133A,C140A) (NMR, 10 STRUCTURES) IPAA 6	
108	1tf3	A	257	314	7.2e-15	0.32	0.36		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	ZINC FINGER TRANSCRIPTION FACTOR SP1; ZINC FINGER, TRANSCRIPTION ACTIVATION, SP1 COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
108	1tf6	A	259	428	1.2e-35			94.31	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
108	1tf6	A	262	404	1.2e-35	0.05	0.99		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
108	1ub6	A	287	426	1.7e-34	-0.10	0.62		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
108	1ubd	C	266	367	1e-32	0.30	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
108	1ubd	C	284	396	8.5e-33			90.80	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
108	1ubd	C	322	424	8.5e-33	-0.06	0.53		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
108	1ubd	C	350	429	1.7e-24	-0.16	0.06		YY1; CHAIN: C; ADENO-	COMPLEX (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
108	2gli	A	259	396	4.8e-66	0.22	1.00		ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
108	2gli	A	260	396	4.8e-66			97.44	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
108	2gli	A	266	394	6.8e-33	0.29	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
108	2gli	A	294	424	8.5e-31	0.26	0.54		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
109	1b7t	A	1	784	0			483.37	MYOSIN HEAVY CHAIN; CHAIN: A; MYOSIN REGULATORY LIGHT CHAIN; CHAIN: Y; MYOSIN ESSENTIAL LIGHT CHAIN; CHAIN: Z;	MYOSIN MYOSIN MOTOR
109	1b7t	A	9	789	0	0.50	1.00		MYOSIN HEAVY CHAIN; CHAIN: A; MYOSIN REGULATORY LIGHT CHAIN; CHAIN: Y;	MYOSIN MYOSIN MOTOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
109	1br1	A	1	740	0	0.61	1.00		MYOSIN ESSENTIAL LIGHT CHAIN; CHAIN: Z	MUSCLE PROTEIN MDE; MUSCLE PROTEIN
109	1br1	A	1	740	0			526.30	MYOSIN; CHAIN: A, B, C, D, E, F, G, H;	MUSCLE PROTEIN MDE; MUSCLE PROTEIN
109	1br2	A	1	710	0	0.65	1.00		MYOSIN; CHAIN: A, B, C, D, E, F, G, H;	MUSCLE PROTEIN MUSCLE PROTEIN
109	1br2	A	1	710	0			489.18	MYOSIN; CHAIN: A, B, C, D, E, F;	MUSCLE PROTEIN MUSCLE PROTEIN
109	1dfk	A	9	789	0	0.33	1.00		MYOSIN HEAD; CHAIN: A; MYOSIN HEAD; CHAIN: Y; MYOSIN HEAD; CHAIN: Z;	CONTRACTILE PROTEIN MYOSIN MOTOR, CONFORMATIONAL CHANGES
109	1lvk		1	710	0			496.50	MYOSIN; CHAIN: NULL;	CONTRACTILE PROTEIN MYOSIN, DICTYOSTELIUM, MOTOR, MANT, ATPASE, ACTIN-BINDING, 2 COILED COIL
109	1lvk		6	710	0	0.48	1.00		MYOSIN; CHAIN: NULL;	CONTRACTILE PROTEIN MYOSIN, DICTYOSTELIUM, MOTOR, MANT, ATPASE, ACTIN-BINDING, 2 COILED COIL
109	1mnd		1	639	0			425.46	MYOSIN; CHAIN: NULL;	CONTRACTILE PROTEIN ATPASE, MYOSIN, COILED COIL, ACTIN-BINDING, ATP BINDING, 2 HEPTAD REPEAT PATTERN, METHYLATION, 3 ALKYLATION, 3 PHOSPHORYLATION, CONTRACTILE PROTEIN
109	1mnd		3	639	0	0.40	1.00		MYOSIN; CHAIN: NULL;	CONTRACTILE PROTEIN ATPASE, MYOSIN, COILED COIL, ACTIN-BINDING, ATP BINDING, 2 HEPTAD REPEAT PATTERN, METHYLATION,

Table 5

SEQ ID NO;	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
109	2mys	A	1	771	0			392.48	MYOSIN; CHAIN: A, B, C;	ALKYLATION, 3 PHOSPHORYLATION, CONTRACTILE PROTEIN c
109	2mys	A	6	743	0	0.58	1.00		MYOSIN; CHAIN: A, B, C;	MUSCLE PROTEIN MUSCLE PROTEIN, MYOSIN SUBFRAGMENT-1, MYOSIN HEAD, 2 MOTOR PROTEIN
110	1a06		10	318	1.7e-89			142.86	CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE; CHAIN: NULL;	KINASE KINASE, SIGNAL TRANSDUCTION, CALCIUM/CALMODULIN
110	1a06		11	300	1.7e-89	0.39	1.00		CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE; CHAIN: NULL;	KINASE KINASE, SIGNAL TRANSDUCTION, CALCIUM/CALMODULIN
110	1a60		1	321	3.6e-56			110.78	PROTEIN KINASE CK2/ALPHA-SUBUNIT; CHAIN: NULL;	TRANSFERASE TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, CASEIN KINASE, 2 SER/THR KINASE
110	1apm	E	17	311	0	0.54	1.00		TRANSFERASE(PHOSPHO TRANSFERASE) \$C-/AMP\$-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (\$C/APKS) 1APM 3 (CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (S139A\$) COMPLEX WITH THE PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 1APM 6	
110	1apm	E	2	326	0			168.62	TRANSFERASE(PHOSPHO TRANSFERASE) \$C-/AMP\$-	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (\$C/APK\$) 1APM 3 (CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (S139A\$) COMPLEX WITH THE PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 1APM 6	
110	1aq1		17	285	3.6e-68			112.67	CYCLIN-DEPENDENT PROTEIN KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; PROTEIN KINASE, CELL CYCLE, PHOSPHORYLATION, STAUROSPORINE, 2 CELL DIVISION, MITOSIS, INHIBITION
110	1aq1		20	271	3.6e-68	0.34	1.00		CYCLIN-DEPENDENT PROTEIN KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; PROTEIN KINASE, CELL CYCLE, PHOSPHORYLATION, STAUROSPORINE, 2 CELL DIVISION, MITOSIS, INHIBITION
110	1cmk	E	16	311	0	0.64	1.00		PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT 1CMK 3 (E.C.2.7.1.37) 1CMK 4	
110	1cmk	E	2	329	0			163.96	PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT 1CMK 3 (E.C.2.7.1.37) 1CMK 4	
110	1csn		18	274	2.4e-66	0.49	1.00		CASEIN KINASE-I; 1CSN 4	PHOSPHOTRANSFERASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
110	1ctp	E	16	304	0	0.77	1.00		TRANSFERASE(PHOSPHO-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (CAK) 1CTP 3 (CATALYTIC SUBUNIT) 1CTP 4	
110	1ctp	E	2	313	0			171.30	TRANSFERASE(PHOSPHO-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (CAK) 1CTP 3 (CATALYTIC SUBUNIT) 1CTP 4	
110	1fgk	A	2	277	3.4e-34			119.24	FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2 PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
110	1fgk	B	1	276	4.8e-47			138.09	FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2 PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
110	1hcl		17	285	8.5e-62			122.03	HUMAN CYCLIN-DEPENDENT KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; TRANSFERASE, SERINE/THREONINE PROTEIN KINASE, ATP-BINDING, 2 CELL CYCLE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
110	ljnk		4	366	3.4e-48			127.21	C-JUN N-TERMINAL KINASE; CHAIN: NULL;	CELL DIVISION, MITOSIS, PHOSPHORYLATION TRANSFERASE JNK3; TRANSFERASE, JNK3 MAP KINASE, SERINE/THREONINE PROTEIN 2 KINASE
110	lkoa		6	275	1.7e-73	0.48	1.00		TWITCHIN; CHAIN: NULL;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
110	lkob	A	3	344	1.2e-92			139.53	TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
110	lkob	A	6	274	1.2e-73	0.39	1.00		TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
110	lkob	A	7	336	1.2e-92	0.36	1.00		TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
110	lp38		5	313	2.4e-63			119.80	MAP KINASE P38; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE; TRANSFERASE, MAP KINASE, SERINE/THREONINE- PROTEIN KINASE, 2 P38
110	lphk		18	272	5.1e-88	0.77	1.00		PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE- PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING
110	lphk		19	275	5.1e-88			170.32	PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE- PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING
110	lki	A	16	331	1.2e-60			124.66	TTIN; CHAIN: A, B;	SERINE KINASE SERINE KINASE, TITIN, MUSCLE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
110	3erk		4	370	1.2e-47			120.67	EXTRACELLULAR REGULATED KINASE 2; CHAIN: NULL;	AUTOINHIBITION TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE, MAP 2, ERK2; TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, MAP KINASE, 2 ERK2
111	1crz	A	176	400	0.00017	0.20	0.31		TOLB PROTEIN; CHAIN: A;	TOXIN BINDING PROTEIN TWO DOMAINS: BETA PROPELLER AND ALPHA/BETA FOLD
111	1erj	A	146	473	3.4e-70	0.49	1.00		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
111	1erj	A	224	514	8.5e-73	0.94	1.00		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
111	1got	B	114	470	1.7e-83			167.73	GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMAL, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2, SIGNAL TRANSDUCTION
111	1got	B	139	377	1.7e-47	0.48	0.98		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMAL, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2, SIGNAL TRANSDUCTION
111	1got	B	156	470	1.7e-83	0.89	1.00		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-	COMPLEX (GTP-BINDING/TRANSDUCER)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
111	1got	B	219	511	5.1e-68	0.79	1.00		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
112	1c4o	A	301	431	3.4e-10	0.58	0.69		DNA NUCLEOTIDE EXCISION REPAIR ENZYME UVRB; CHAIN: A;	REPLICATION DNA NUCLEOTIDE EXCISION REPAIR, UVRABC, HELICASE, 2 HYPERTHERMOSTABLE PROTEIN
112	1d2m	A	301	431	3.4e-10	0.52	0.25		EXCINUCLEASE ABC SUBUNIT B; CHAIN: A;	HYDROLASE UVRB; MULTIDOMAIN PROTEIN
112	1d9x	A	301	468	1e-14	0.07	0.35		EXCINUCLEASE UVRABC COMPONENT UVRB; CHAIN: A;	GENE REGULATION APO PROTEIN
112	1fik	A	285	445	5.1e-36	0.10	-0.05		EUKARYOTIC INITIATION FACTOR 4A; CHAIN: A;	TRANSLATION YEAST INITIATION FACTOR 4A, EIF4A; HELICASE, INITIATION FACTOR 4A, DEAD-BOX PROTEIN
112	1fuu	B	2	120	0.0024	-0.10	0.06		YEAST INITIATION FACTOR 4A; CHAIN: A, B;	TRANSLATION EUKARYOTIC INITIATION FACTOR 4A; IF4A, HELICASE, DEAD-BOX PROTEIN
112	1fuu	B	56	445	1.2e-91	-0.31	0.04		YEAST INITIATION	TRANSLATION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
112	1qun	B	1227	1347	1.2e-08	0.14	-0.20		FACTOR 4A; CHAIN: A, B; EUKARYOTIC INITIATION FACTOR 4A; IF4A, HELICASE, DEAD-BOX PROTEIN	
112	1qun	B	1241	1307	9.6e-10	0.29	-0.20		PAPD-LIKE CHAPERONE FIMC; CHAIN: A, C, E, G, I, K, M, O; MANNOSE-SPECIFIC ADHESIN FIMH; CHAIN: B, D, F, H, J, L, N, P;	CHAPERONE/STRUCTURAL PROTEIN CHAPERONE ADHESIN DONOR STRAND COMPLEMENTATION, 2 CHAPERONE/STRUCTURAL PROTEIN
114	1asy		89	293	1.7e-35	0.25	0.03		PROTEIN TYROSINE PHOSPHATASE 1B; CHAIN: NULL;	HYDROLASE HYDROLASE, DEPHOSPHORYLATION
114	1c83	A	89	293	1.7e-38	0.20	-0.09		PROTEIN-TYROSINE PHOSPHATASE 1B; CHAIN: A;	HYDROLASE PTP1B; HYDROLASE, PHOSPHORYLATION, LIGAND, INHIBITOR
114	1d5r	A	122	302	3.4e-20	0.38	0.95		PHOSPHOINOSITIDE PHOSPHOTASE PTEN; CHAIN: A;	HYDROLASE C2 DOMAIN, PHOSPHOTIDYLINOSITOL, PHOSPHOTASE, HYDROLASE
114	1gwz		89	301	8.5e-40	0.18	0.01		SHP-1; CHAIN: NULL;	HYDROLASE PROTEIN-TYROSINE PHOSPHATASE; HYDROLASE, PROTEIN
114	1mkp		159	296	2.4e-20	0.14	0.71		PYST1; CHAIN: NULL;	TYROSINE PHOSPHATASE, CATALYTIC DOMAIN, 2 WPD LOOP, SH2 DOMAIN
114	1mkp		185	296	1.7e-12	-0.04	0.48		PYST1; CHAIN: NULL;	HYDROLASE DUAL SPECIFICITY PHOSPHATASE, MAP KINASE HYDROLASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
114	1rpm	A	84	298	1.5e-43	-0.10	0.09		RECEPTOR PROTEIN TYROSINE PHOSPHATASE MU; CHAIN: A, B;	SPECIFICITY PHOSPHATASE, MAP KINASE HYDROLASE
114	1vhr	A	124	296	2.4e-20	0.49	0.99		HUMAN VHI-RELATED DUAL-SPECIFICITY PHOSPHATASE CHAIN: A, B;	RECEPTOR D1; RECEPTOR, PHOSPHATASE, SIGNAL TRANSDUCTION, ADHESION, 2 HYDROLASE
114	1vhr	A	160	301	6.8e-11	0.38	0.40		HUMAN VHI-RELATED DUAL-SPECIFICITY PHOSPHATASE CHAIN: A, B;	HYDROLASE VHR; HYDROLASE, PROTEIN DUAL-SPECIFICITY PHOSPHATASE
114	1yfo	A	62	298	6.8e-46	0.31	0.09		RECEPTOR PROTEIN TYROSINE PHOSPHATASE ALPHA; CHAIN: A, B;	HYDROLASE D1; HYDROLASE, SIGNAL TRANSDUCTION, RECEPTOR, GLYCOPROTEIN, 2 PHOSPHORYLATION, SIGNAL
114	1ym		115	289	6.8e-12	-0.05	0.00		YERSINIA PROTEIN TYROSINE PHOSPHATASE; CHAIN: NULL;	HYDROLASE YOP51, YOP2B, PASTEURELLA X, PTP-ASE, PROTEIN TYROSINE PHOSPHATASE, HYDROLASE
114	1ym		205	294	3.6e-05	-0.07	0.39		YERSINIA PROTEIN TYROSINE PHOSPHATASE; CHAIN: NULL;	HYDROLASE YOP51, YOP2B, PASTEURELLA X, PTP-ASE, PROTEIN TYROSINE PHOSPHATASE, HYDROLASE
114	2shp	A	59	298	1.4e-48	0.17	-0.12		SHP-2; CHAIN: A, B;	TYROSINE PHOSPHATASE SYP, SHPTP-2; TYROSINE PHOSPHATASE, INSULIN SIGNALING, SH2 PROTEIN
115	1euv	A	364	588	5.1e-41	0.51	1.00		ULP1 PROTEASE; CHAIN: A; UBIQUITIN-LIKE	HYDROLASE SUMO HYDROLASE, UBIQUITIN-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									PROTEIN SMT3; CHAIN: B;	LIKE PROTEASE 1, SMT3 HYDROLASE 2 DESUMOYLATING ENZYME, CYSTEINE PROTEASE, SUMO PROCESSING 3 ENZYME, SMT3 PROCESSING ENZYME, NABH4, THIOHEMIACETAL, 4 COVALENT PROTEASE ADDUCT
118	1ez3	A	145	265	6e-11	0.29	-0.17		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
118	1ez3	A	174	293	4.8e-09	0.28	-0.17		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
118	1ez3	A	206	327	1.2e-11	0.15	-0.06		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
118	1ez3	A	307	425	2.4e-11	0.15	-0.18		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
118	1ez3	A	378	499	1.1e-10	0.25	-0.19		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
118	1ez3	A	52	151	8.4e-10	0.11	-0.18		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
118	1ez3	A	93	224	8.4e-13	0.08	-0.18		SYNTAXIN-1A; CHAIN: A, B, C;	ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
118	1quu	A	736	971	6e-16	0.02	-0.12		HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
118	1sig		52	251	1.2e-09	-0.24	0.06		RNA POLYMERASE PRIMARY SIGMA FACTOR; CHAIN: NULL;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
120	1amp		193	356	1.2e-18	0.04	0.43		HYDROLASE(AMINOPEPTIDASE) AMINOPEPTIDASE (AEROMONAS PROTEOLYTICA) (E.C.3.4.11.10) IAMP 3	TRANSCRIPTION REGULATION SIGMA70; RNA POLYMERASE SIGMA FACTOR, TRANSCRIPTION REGULATION
120	1amp		211	468	5.1e-30	0.11	0.43		HYDROLASE(AMINOPEPTIDASE) AMINOPEPTIDASE (AEROMONAS PROTEOLYTICA) (E.C.3.4.11.10) IAMP 3	
120	1de4	C	224	592	1e-46	0.02	0.83		HEMOCHROMATOSIS PROTEIN; CHAIN: A, D, G; BETA-2-MICROGLOBULIN; CHAIN: B, E, H; TRANSFERRIN RECEPTOR; CHAIN: C, F, I; AMINOPEPTIDASE; CHAIN: A;	METAL TRANSPORT INHIBITOR/RECEPTOR HFE; HFE, HEREDITARY HEMOCHROMATOSIS, MHC CLASS I, TRANSFERRIN 2 RECEPTOR
120	1qq9	A	211	463	1.7e-28	-0.00	0.64			HYDROLASE SGAP; DOUBLE-ZINC METALLOPROTEINASE, CALCIUM ACTIVATION, PROTEIN-2 INHIBITOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										COMPLEX
121	1cfl	A	1	358	1.2e-41			73.18	ARRESTIN; CHAIN: A, B, C, D;	STRUCTURAL PROTEIN RETINAL S-ANTIGEN, 48 KD PROTEIN; VISUAL ARRESTIN, DESENSITISATION OF THE VISUAL TRANSDUCTION 2 CASCADE, BINDING TO ACTICATED AND PHOSPHORYLATED RHODOPSIN
121	1cfl	A	65	265	1.2e-41	-0.35	0.00		ARRESTIN; CHAIN: A, B, C, D;	STRUCTURAL PROTEIN RETINAL S-ANTIGEN, 48 KD PROTEIN; VISUAL ARRESTIN, DESENSITISATION OF THE VISUAL TRANSDUCTION 2 CASCADE, BINDING TO ACTICATED AND PHOSPHORYLATED RHODOPSIN
121	1cfl	D	1	351	1.2e-43			71.95	ARRESTIN; CHAIN: A, B, C, D;	STRUCTURAL PROTEIN RETINAL S-ANTIGEN, 48 KD PROTEIN; VISUAL ARRESTIN, DESENSITISATION OF THE VISUAL TRANSDUCTION 2 CASCADE, BINDING TO ACTICATED AND PHOSPHORYLATED RHODOPSIN
122	1pma	B	12	205	1.2e-44			71.75	PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	PROTEASE PROSOME, MULTICATALYTIC PROTEASE, MCP, MACROPAIN; PROTEASE, PROTEASOME, HYDROLASE
122	1ryp	H	1	205	1.4e-43			55.61	20S PROTEASOME;	MULTICATALYTIC

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
122	1ryp	J	1	205	1.5e-36			84.34	20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
122	1ryp	K	1	205	1.5e-36			58.38	20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
122	1ryp	L	2	205	3.4e-48			52.75	20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
124	1div		94	240	6.8e-42	0.38	0.82		RIBOSOMAL PROTEIN L9; CHAIN: NULL;	RIBOSOMAL PROTEIN RIBOSOMAL PROTEIN, RNA-BINDING
124	1div		94	243	6.8e-42			54.38	RIBOSOMAL PROTEIN L9; CHAIN: NULL;	RIBOSOMAL PROTEIN RIBOSOMAL PROTEIN, RNA-BINDING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
128	1a88	A	74	366	1.4e-46			68.67	CHLOROPEROXIDASE L; CHAIN: A, B, C;	HALOPEROXIDASE BROMOPEROXIDASE L; HALOPEROXIDASE L; HALOPEROXIDASE, OXIDOREDUCTASE
128	1a8q		77	366	1.7e-40			54.47	BROMOPEROXIDASE A1; CHAIN: NULL;	HALOPEROXIDASE CHLOROPEROXIDASE A1, HALOPEROXIDASE A1; HALOPEROXIDASE, OXIDOREDUCTASE
128	1a8s		72	366	1.7e-44			50.66	CHLOROPEROXIDASE F; CHAIN: NULL;	HALOPEROXIDASE HALOPEROXIDASE F; HALOPEROXIDASE, OXIDOREDUCTASE, PROIONATE COMPLEX
128	1azw	A	59	367	5.1e-37			56.93	PROLINE IMINOPEPTIDASE; CHAIN: A, B;	AMINOPEPTIDASE AMINOPEPTIDASE, PROLINE IMINOPEPTIDASE, SERINE PROTEASE, 2 XANTHOMONAS CAMPESTRIS
128	1b6g		51	353	1.2e-31			71.01	HALOALKANE DEHALOGENASE; CHAIN: NULL;	HYDROLASE HYDROLASE, HALOALKANE DEHALOGENASE, ALPHA/BETA-HYDROLASE
128	1brt		72	365	1.7e-38			68.50	BROMOPEROXIDASE A2; CHAIN: NULL;	HALOPEROXIDASE HALOPEROXIDASE A2, CHLOROPEROXIDASE A2; HALOPEROXIDASE, OXIDOREDUCTASE, PEROXIDASE, ALPHA/BETA 2 HYDROLASE FOLD, MUTANT M99T
128	1c4x	A	67	366	1.4e-40			63.30	2-HYDROXY-6-OXO-6- PHENYLHEXA-2,4- DIENOATE CHAIN: A;	HYDROLASE BPHD; HYDROLASE, PCB DEGRADATION
128	1cqw	A	65	368	6.8e-43			73.30	HALOALKANE DEHALOGENASE; 1- CHLOROHEXANE CHAIN:	HYDROLASE A/B HYDROLASE FOLD, DEHALOGENASE I-S BOND

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
128	1ehy	A	74	365	1.7e-45			70.69	A; SOLUBLE EPOXIDE HYDROLASE; CHAIN: A, B, C, D;	HYDROLASE HYDROLASE, ALPHA/BETA HYDROLASE FOLD, EPOXIDE DEGRADATION, 2 EPICHLOROHYDRIN
128	1jkm	A	1	342	1.7e-06			52.92	BREFELDIN A ESTERASE; CHAIN: A, B;	SERINE HYDROLASE, SERINE HYDROLASE, DEGRADATION OF BREFELDIN A, ALPHA/BETA 2 HYDROLASE FAMILY
128	1qj4	A	93	368	8.5e-17			50.42	HYDROXYNITRILE LYASE; CHAIN: A;	LYASE OXYNITRILE LYASE; OXYNITRILASE, CYANOGENESIS, CYANHYDRIN FORMATION, LYASE
128	1qtr	A	59	366	3.4e-35			75.39	PROLYL AMINOPEPTIDASE; CHAIN: A;	HYDROLASE ALPHA BETA HYDROLASE FOLD, PROLINE, PROLYL AMINOPEPTIDASE, 2 SERRATIA, IMINOPEPTIDASE
133	1amu	A	84	654	0			159.04	GRAMICIDIN SYNTHETASE 1; CHAIN: A, B; PHENYLALANINE; CHAIN: C, D;	PEPTIDE SYNTHETASE GRSA; PEPTIDE SYNTHETASE, GRSA, ADENYLATE FORMING
133	1lci		93	650	0			192.59	LUCIFERASE; CHAIN: NULL;	OXIDOREDUCTASE OXIDOREDUCTASE, MONOOXYGENASE, PHOTOPROTEIN, LUMINESCENCE
138	1tub	A	1	232	0			260.21	TUBULIN; CHAIN: A, B;	MICROTUBULES, ALPHA-TUBULIN, BETA-TUBULIN, GTPASE HELIX
141	1klo		189	350	3.4e-20			115.60	LAMININ; CHAIN: NULL;	GLYCOPROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										GLYCOPROTEIN
143	1alh	A	531	613	1e-29			81.72	QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
143	1buo	A	19	143	8.5e-19			77.84	PROMYELOCYTIC LEUKEMIA ZINC FINGER PROTEIN PLZF; CHAIN: A;	GENE REGULATION POZ DOMAIN; PROTEIN-PROTEIN INTERACTION DOMAIN, TRANSCRIPTIONAL 2 REPRESSOR, ZINC-FINGER PROTEIN, X-RAY CRYSTALLOGRAPHY, 3 PROTEIN STRUCTURE, PROMYELOCYTIC LEUKEMIA, GENE REGULATION
143	1mey	C	530	612	3.4e-48			84.00	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
143	1tf6	A	437	609	5.1e-30			79.59	TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE II, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
143	1ubd	C	504	612	1.7e-32			86.70	YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
143	2gli	A	473	613	1.7e-31			71.69	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	ELEMENT, YY1, ZINC 2 FINGER PROTEIN; DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
154	1dhp	A	1	272	3.4e-83			113.45	DIHYDRODIPICOLINATE SYNTHASE; CHAIN: A, B;	SYNTHASE DHDPs; SYNTHASE, DIHYDRODIPICOLINATE
154	1nal	1	1	268	1.2e-55			97.53	N-ACETYLNEURAMINATE LYASE; INAL 4 CHAIN: 1, 2, 3, 4; INAL 5	LYASE
157	1a06		209	510	8.5e-60			98.67	CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE; CHAIN: NULL;	KINASE KINASE, SIGNAL TRANSDUCTION, CALCIUM/CALMODULIN
157	1a60		187	83	1e-86			129.78	PROTEIN KINASE CK2/ALPHA-SUBUNIT; CHAIN: NULL;	TRANSFERASE, TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, CASEIN KINASE, 2 SER/THR KINASE
157	1apm	E	183	533	8.5e-48			115.15	TRANSFERASE(PHOSPHO TRANSFERASE) \$C-/AMP\$-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (SC/APK\$) 1APM 3 (CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (/S139A\$) COMPLEX WITH THE PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
157	1aq1		214	524	0			160.34	MEGA-8 1APM 6 CYCLIN-DEPENDENT PROTEIN KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; PROTEIN KINASE, CELL CYCLE, PHOSPHORYLATION, STAUROSPORINE, 2 CELL DIVISION, MITOSIS, INHIBITION
157	1bi8	A	214	513	1.2e-91			148.04	CYCLIN-DEPENDENT KINASE 6; CHAIN: A, C; CYCLIN-DEPENDENT KINASE INHIBITOR; CHAIN: B, D;	COMPLEX (KINASE/INHIBITOR) CDK6; P19INK4D; CYCLIN DEPENDENT KINASE, CYCLIN DEPENDENT KINASE INHIBITORY 2 PROTEIN, CDK, INK4, CELL CYCLE, COMPLEX (KINASE/INHIBITOR) HEADER HELIX
157	1blx	A	210	83	6.8e-100			163.82	CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
157	1cmk	E	163	533	5.1e-49			116.16	PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT 1CMK 3 (E.C.2.7.1.37) 1CMK 4	
157	1ctp	E	181	509	6.8e-50			112.97	TRANSFERASE(PHOSPHO TRANSFERASE) CAMP- DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (CAPK) 1CTP 3 (CATALYTIC SUBUNIT) 1CTP 4	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
157	1hcl		214	524	0			171.50	HUMAN CYCLIN-DEPENDENT KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; TRANSFERASE, SERINE/THREONINE PROTEIN KINASE, ATP-BINDING, 2 CELL CYCLE, CELL DIVISION, MITOSIS, PHOSPHORYLATION
157	1ian		200	120	0			145.96	P38 MAP KINASE; CHAIN: NULL;	SERINE/THREONINE-PROTEIN KINASE CSBP, RK, P38; PROTEIN SER/THR-KINASE, SERINE/THREONINE-PROTEIN KINASE
157	1jnk		200	544	0			162.87	C-JUN N-TERMINAL KINASE; CHAIN: NULL;	TRANSFERASE JNK3; TRANSFERASE, JNK3 MAP KINASE, SERINE/THREONINE PROTEIN 2 KINASE
157	1koa		192	629	8.5e-57			133.78	TWITCHIN; CHAIN: NULL;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
157	1kob	A	187	540	8.5e-54			134.02	TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
157	1p38		199	560	0			174.44	MAP KINASE P38; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE; TRANSFERASE, MAP KINASE, SERINE/THREONINE-PROTEIN KINASE, 2 P38
157	1phk		214	482	6.8e-57			96.25	PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE-PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING
157	1pme		210	557	0			161.09	ERK2; CHAIN: NULL;	TRANSFERASE MAP KINASE, SERINE/THREONINE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
157	1tki	A	210	534	1.4e-45			127.01	TITIN; CHAIN: A, B;	PROTEIN KINASE, TRANSFERASE SERINE KINASE SERINE KINASE, TITIN, MUSCLE, AUTONHIBITION
157	3erk		198	553	0			165.77	EXTRACELLULAR REGULATED KINASE 2; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE, MAP 2, ERK2; TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, MAP KINASE, 2 ERK2
163	1a9n	A	136	304	3.4e-28			54.53	U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A; CHAIN: A, C; U2 B; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
163	1a9n	C	136	317	1.2e-28			53.48	U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A; CHAIN: A, C; U2 B; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
168	1alh	A	388	470	5.1e-30			74.63	QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
168	1mey	C	247	329	1.4e-48			97.41	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
168	1tf6	A	216	384	5.1e-37			104.01	TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
168	1ubd	C	473	581	3.4e-35			91.10	YY1; CHAIN: C; ADENOVIRUS ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
168	2gli	A	222	358	1.2e-34			99.34	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
171	1a17		274	430	5.1e-22			96.15	SERINE/THREONINE PROTEIN PHOSPHATASE 5; CHAIN: NULL;	HYDROLASE TETRATRICOPEPTIDE, TRP; HYDROLASE, PROTEIN-PHOSPHATASE, PROTEIN-PROTEIN INTERACTIONS, TPR, 2 SUPER-HELIX, X-RAY STRUCTURE
173	1by1	A	150	362	1.2e-30			99.42	PIX; CHAIN: A;	TRANSPORT PROTEIN RHO-GTPASE EXCHANGE FACTOR, TRANSPORT PROTEIN
173	1fmk		70	495	1.7e-100			73.30	TYROSINE-PROTEIN KINASE SRC; CHAIN: NULL;	PHOSPHOTRANSFERASE C-SRC, P60-SRC; SRC, TYROSINE KINASE, PHOSPHORYLATION, SH2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
173	1qcf	A	68	490	3.4e-100			74.45	HAEMATOPOIETIC CELL KINASE (HCK); CHAIN: A;	SH3, 2 PHOSPHOTYROSINE, PROTO-ONCOGENE, PHOSPHOTRANSFERASE
186	1mev	C	103	185	5.1e-50			73.78	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	TYROSINE KINASE TYROSINE KINASE-INHIBITOR COMPLEX, DOWN-REGULATED KINASE, 2 ORDERED ACTIVATION LOOP
188	1a79	A	1	172	1.7e-42			61.23	TRNA ENDONUCLEASE; CHAIN: A, B, C, D;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
191	1a06		139	458	3.4e-81			119.44	CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE; CHAIN: NULL;	ENDONUCLEASE ENDONUCLEASE, TRNA ENDONUCLEASE
191	1apm	E	113	478	0			251.49	TRANSFERASE (PHOSPHOTRANSFERASE) \$C\text{-}AMP\text{-}DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (\$C/APK\$) 1APM 3 (CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (S139A\$) COMPLEX WITH THB PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 1APM 6	KINASE KINASE, SIGNAL TRANSDUCTION, CALCIUM/CALMODULIN

Table 5

SEQ ID NO;	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
191	1aq1		145	432	5.1e-51			110.89	CYCLIN-DEPENDENT PROTEIN KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; PROTEIN KINASE, CELL CYCLE, PHOSPHORYLATION, STAUROSPORINE, 2 CELL DIVISION, MITOSIS, INHIBITION
191	1bi8	A	146	458	3.4e-39			112.00	CYCLIN-DEPENDENT KINASE 6; CHAIN: A, C; CYCLIN-DEPENDENT KINASE INHIBITOR; CHAIN: B, D;	COMPLEX (KINASE/INHIBITOR) CDK6; P19INK4D; CYCLIN DEPENDENT KINASE, CYCLIN DEPENDENT KINASE INHIBITORY 2 PROTEIN, CDK, INK4, CELL CYCLE, COMPLEX (KINASE/INHIBITOR) HEADER HELIX
191	1blx	A	140	465	1.5e-42			135.09	CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
191	1cmk	E	105	478	0			252.68	PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT 1CMK 3 (E.C.2.7.1.37) 1CMK 4	
191	1ctp	E	110	459	0			244.28	TRANSFERASE(PHOSPHOTRANSFERASE) CAMP-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (CAPK) 1CTP 3 (CATALYTIC SUBUNIT) 1CTP 4	
191	1hcl		145	432	1.7e-53			123.56	HUMAN CYCLIN-	PROTEIN KINASE CDK2;

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									DEPENDENT KINASE 2; CHAIN: NULL;	TRANSFERASE, SERINE/THREONINE PROTEIN KINASE, ATP-BINDING, 2 CELL CYCLE, CELL DIVISION, MITOSIS, PHOSPHORYLATION
191	1kob	A	118	488	5.1e-64			134.48	TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
191	1p38		128	508	3.4e-46			119.00	MAP KINASE P38; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE; TRANSFERASE, MAP KINASE, SERINE/THREONINE-PROTEIN KINASE, 2 P38
191	1phk		144	424	1.7e-74			132.88	PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE-PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING
191	1tki	A	142	475	3.4e-52			114.63	TITIN; CHAIN: A, B;	SERINE KINASE SERINE KINASE, TITIN, MUSCLE, AUTOINHIBITION
191	3erk		132	506	3.4e-43			121.20	EXTRACELLULAR REGULATED KINASE 2; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE, MAP 2, ERK2; TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, MAP KINASE, 2 ERK2
193	1a9w	E	2	78	6.8e-30			76.31	HEMOGLOBIN; CHAIN: A, E, C, F;	OXYGEN TRANSPORT
193	1fdh	G	3	78	3.4e-29			67.75	OXYGEN TRANSPORT HEMOGLOBIN (DEOXY, HUMAN FETAL F ₂ -H ₂) 1FDHG 1 1FDHH 2	OXYGEN TRANSPORT

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
195	1got	B	1	311	1.5e-45			66.17	GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
199	lawc	B	2	169	1.4e-39			94.15	GA BINDING PROTEIN ALPHA; CHAIN: A; GA BINDING PROTEIN BETA 1; CHAIN: B; DNA; CHAIN: D, E;	COMPLEX (TRANSCRIPTION REGULATION/DNA) GABPALPHA; GABPBETA1; COMPLEX (TRANSCRIPTION REGULATION/DNA), DNA-BINDING, 2 NUCLEAR PROTEIN, ETS DOMAIN, ANKYRIN REPEATS, TRANSCRIPTION 3 FACTOR
199	likn	D	48	262	3.4e-40			83.26	NF-KAPPA-B P65 SUBUNIT; CHAIN: A; NF-KAPPA-B P50D SUBUNIT; CHAIN: C; I-KAPPA-B-ALPHA; CHAIN: D;	TRANSCRIPTION FACTOR P65; P50D; TRANSCRIPTION FACTOR, IKB/NFKB COMPLEX
199	lnfi	E	38	249	6.8e-41			91.60	NF-KAPPA-B P65; CHAIN: A, C; NF-KAPPA-B P50; CHAIN: B, D; I-KAPPA-B-ALPHA; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REG/ANK REPEAT) COMPLEX (TRANSCRIPTION REGULATION/ANK REPEAT), ANKYRIN 2 REPEAT HELIX
217	laog	A	149	597	1e-58			78.99	TRYPTOPHAN THIONE REDUCTASE; CHAIN: A, B;	OXIDOREDUCTASE TRYPTOPHAN THIONE REDUCTASE, FAD DEPENDENT DISULPHIDE 2
217	lebd	A	157	597	1.5e-91			115.94	DIHYDROLIPOAMIDE DEHYDROGENASE;	OXIDOREDUCTASE COMPLEX (OXIDOREDUCTASE/TRANS

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
217	1fed	A	193	578	0.00051			77.29	CHAIN: A, B; DIHYDROLIPOAMIDE ACETYLTRANSFERASE; CHAIN: C; ELECTRON TRANSPORT(FLAVOCYTO CHROME) FLAVOCYTOCHROME C SULFIDE DEHYDROGENASE (FCSO) IFCD 3	FERASE) E3BD; REDOX- ACTIVE CENTER, GLYCOLYSIS, OXIDOREDUCTASE
217	1fec	A	132	597	1.7e-58			91.61	TRYPTANTHIONE REDUCTASE; CHAIN: A, B;	OXIDOREDUCTASE REDOX- ACTIVE CENTER, OXIDOREDUCTASE, FLAVOPROTEIN, FAD, NADP
217	1ges	A	160	593	1.4e-77			100.73	OXIDOREDUCTASE(FLAV OENZYME) GLUTATHIONE REDUCTASE (E.C.1.6.4.2) NAD MUTANT WITH ALA 179 1GES 3 REPLACED BY GLY, ALA 183 BY GLY, VAL 197 BY GLU, ARG 198 BY 1GES 4 MET, LYS 199 BY PHE, HIS 200 BY ASP, AND ARG 204 BY PHE 1GES 5 (A179G,A183G,V197E,R198 M,K199F,H200D,R204P) COMPLEXED WITH 1GES 6 NAD 1GES 7	
217	1lpf	A	151	597	3.4e-83			112.80	OXIDOREDUCTASE DIHYDROLIPOAMIDE DEHYDROGENASE (E.C.1.8.1.4) COMPLEX WITH 1LPF 3 FLAVIN- ADENINE-DINUCLEOTIDE (FAD) 1LPF 4	
217	1lvi		148	597	5.1e-70			89.94	OXIDOREDUCTASE	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Pst Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
217	1nhp		195	597	3.4e-64			157.16	DIHYDROLIPOAMIDE DEHYDROGENASE (E.C.1.8.1.4) COMPLEX WITH 1LVL 3 NICOTINAMIDE-ADENINE-DINUCLEOTIDE (NAD+) 1LVL 4	
217	1ojt		147	597	5.1e-76			81.98	OXIDOREDUCTASE (H2O2(A)) NADH PEROXIDASE (NPX) (E.C.1.11.1.1) MUTANT WITH CYS 42 INHP 3 REPLACED BY ALA (C42A) INHP 4	OXIDOREDUCTASE REDOX-ACTIVE CENTER, GLYCOLYSIS, OXIDOREDUCTASE, NAD, 2 FLAVOPROTEIN, FAD, F64K
217	1tb		190	490	1.2e-40			91.34	OXIDOREDUCTASE(FLAVOENZYME) THIOREDOXIN REDUCTASE (E.C.1.6.4.5) MUTANT WITH CYS 138 ITRB 3 REPLACED BY SER (C138S) ITRB 4	
217	1vdc		174	495	5.1e-39			92.67	NADPH DEPENDENT THIOREDOXIN REDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE NTR; HYPOTHETICAL PROTEIN, REDOX-ACTIVE CENTER, OXIDOREDUCTASE, 2 DISULFIDE OXIDOREDUCTASE, THIOREDOXIN REDUCTASE, FLAVIN 3 ADENINE DINUCLEOTIDE
217	3grs		158	593	8.5e-71			97.61	OXIDOREDUCTASE (FLAVOENZYME) GLUTATHIONE REDUCTASE (E.C.1.6.4.2), OXIDIZED FORM (B) 3GRS	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
217	3lad	A	151	597	3.4e-82			114.03	OXIDOREDUCTASE DIHYDROLIPOAMIDE DEHYDROGENASE (E.C.1.8.1.4) 3LAD 3	
219	1a06		1	245	8.5e-76			106.55	CALCIUM/CALMODULIN- DEPENDENT PROTEIN KINASE; CHAIN: NULL;	KINASE KINASE, SIGNAL TRANSDUCTION, CALCIUM/CALMODULIN
219	1apm	E	1	254	6.8e-95			104.03	TRANSFERASE(PHOSPHO TRANSFERASE) \$C\$-AMP\$- DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (\$C/APK\$) 1APM 3 (CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (S139AS) COMPLEX WITH THE PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 1APM 6	
219	1bi8	A	1	235	3.4e-38			72.81	CYCLIN-DEPENDENT KINASE 6; CHAIN: A, C; CYCLIN-DEPENDENT KINASE INHIBITOR; CHAIN: B, D;	COMPLEX (KINASE/INHIBITOR) CDK6; P19INK4D; CYCLIN DEPENDENT KINASE, CYCLIN DEPENDENT KINASE INHIBITORY 2 PROTEIN, CDK, INK4, CELL CYCLE, COMPLEX (KINASE/INHIBITOR) HEADER HELIX
219	1cmk	E	1	254	8.5e-96			100.79	PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT 1CMK 3 (E.C.2.7.1.37) 1CMK 4	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
219	1ctp	B	1	250	5.1e-91			100.24	TRANSFERASE(PHOSPHO TRANSFERASE) CAMP- DEPENDENT PROTEIN KINASE (B.C.2.7.1.37) (CAPK) 1CTP 3 (CATALYTIC SUBUNIT) 1CTP 4	
219	1hcl		1	247	3.4e-48			78.81	HUMAN CYCLIN- DEPENDENT KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; TRANSFERASE, SERINE/THREONINE PROTEIN KINASE, ATP- BINDING, 2 CELL CYCLE, CELL DIVISION, MITOSIS, PHOSPHORYLATION
219	1kob	A	1	280	3.4e-59			104.78	TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
219	1phk		1	208	1.4e-73			110.63	PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE- PROTEIN, 2 KINASE, ATP- BINDING, CALMODULIN- BINDING
219	1pme		1	270	1e-41			73.28	ERK2; CHAIN: NULL;	TRANSFERASE MAP KINASE, SERINE/THREONINE PROTEIN KINASE, TRANSFERASE
219	1tki	A	1	255	6.8e-48			75.92	TTTN; CHAIN: A, B;	SERINE KINASE SERINE KINASE, TITIN, MUSCLE, AUTOINHIBITION
222	1cii		167	755	6.8e-05			158.52	COLICIN IA; CHAIN: NULL;	TRANSMEMBRANE PROTEIN COLICIN, BACTERIOCIN, ION CHANNEL FORMATION, TRANSMEMBRANE 2 PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
231	1mkn	A	76	139	8.5e-22			68.94	MIDKINE; CHAIN: A;	HEPARIN-BINDING GROWTH FACTOR HEPARIN-BINDING GROWTH FACTOR
236	lrgs		454	724	1.5e-50			84.25	CAMP DEPENDENT PROTEIN KINASE; CHAIN: NULL;	KINASE R(1ALPHA); REGULATORY SUBUNIT, KINASE
241	lbo1	A	39	421	0			354.11	PHOSPHATIDYLINOSITOL PHOSPHATE KINASE IIBETA; CHAIN: A, B;	TRANSFERASE PPK; LIPID SIGNALING, TRANSFERASE
242	ledh	A	1602	1804	1.5e-20			101.31	E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
242	lncj	A	1609	1811	6.8e-25			106.57	N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
244	lbih	A	1	398	5.1e-27			80.44	HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
244	lcfb		166	345	8.5e-12			73.85	NEURAL ADHESION MOLECULE DROSOPHILA NEUROGLIAN (CHYMOTRYPTIC FRAGMENT CONTAINING THE 1CFB 3 TWO AMINO PROXIMAL FIBRONECTIN TYPE III REPEATS 1CFB 4 (RESIDUES 610 - 814)) 1CFB 5	
244	lfnf		57	435	8.5e-42			136.59	FIBRONECTIN; 1FNF 6	CELL ADHESION PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									CHAIN: NULL; 1FNF 7.	RGD, EXTRACELLULAR MATRIX 1FNF 18
244	1fhh	A	163	435	5.1e-30			83.24	FIBRONECTIN; CHAIN: A;	HEPARIN AND INTEGRIN BINDING HEPARIN AND INTEGRIN BINDING
244	1qg3	A	57	247	3.4e-20			91.86	INTEGRIN BETA-4 SUBUNIT; CHAIN: A, B;	STRUCTURAL PROTEIN INTEGRIN, HEMIDESMOSOME, FIBRONECTIN, CARCINOMA, STRUCTURAL 2 PROTEIN
245	1ayz	A	97	258	5.1e-33			94.52	UBIQUITIN-CONJUGATING ENZYME RAD6; CHAIN: A, B, C;	UBIQUITIN CONJUGATION UBC2; UBIQUITIN CONJUGATION, UBIQUITIN-CONJUGATING ENZYME
245	1c4z	D	103	251	1.2e-24			71.13	UBIQUITIN-PROTEIN LIGASE E3A; CHAIN: A, B, C; UBIQUITIN CONJUGATING ENZYME E2; CHAIN: D;	LIGASE E6AP; UBC7; BILBOAL STRUCTURE, ELONGATED SHAPE, E3 UBIQUITIN LIGASE, E2 2 UBIQUITIN CONJUGATING ENZYME
245	1qcq	A	99	253	3.4e-36			93.22	UBIQUITIN CONJUGATING ENZYME; CHAIN: A;	LIGASE UBIQUITIN, UBIQUITIN-CONJUGATING ENZYME, YEAST
245	1u9a	A	95	252	3.4e-28			83.68	UBC9; CHAIN: NULL;	UBIQUITIN-CONJUGATING ENZYME UBIQUITIN-CONJUGATING ENZYME; UBIQUITIN-CONJUGATING ENZYME, UBIQUITIN-DIRECTED 2 PROTEOLYSIS, CELL CYCLE CONTROL, LIGASE
245	2aak		96	250	6.8e-35			96.93	UBIQUITIN CONJUGATING ENZYME; CHAIN: NULL;	UBIQUITIN CONJUGATION UBC1; UBIQUITIN CONJUGATION, LIGASE
245	2e2c		91	256	6.8e-30			84.68	UBIQUITIN CONJUGATING ENZYME;	UBIQUITIN CONJUGATION, UBIQUITIN CONJUGATING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
245	2ucz		97	250	1.2e-29			88.78	CHAIN: NULL; UBIQUITIN CONJUGATING ENZYME; CHAIN: NULL;	UBIQUITIN CARRIER PROTEIN, THIOESTER 2 BOND, LIGASE UBIQUITIN CONJUGATION UBC7; UBIQUITIN CONJUGATION, LIGASE, YEAST
247	1adl		162	294	8.5e-54			83.63	LIPID-BINDING PROTEIN ADIPOCYTE LIPID- BINDING PROTEIN COMPLEXED WITH ARACHIDONIC 1ADL 3 ACID 1ADL 4	
247	1hmr		162	294	5.1e-56			90.26	LIPID-BINDING PROTEIN FATTY ACID BINDING PROTEIN (HUMAN MUSCLE, M-FABP) COMPLEXED IHMR 3 WITH ONE MOLECULE OF ELADIC ACID IHMR 4	
247	1opa	A	162	294	8.5e-49			198.37	RETINOL TRANSPORT CELLULAR RETINOL BINDING PROTEIN II (APO FORM) (APO-CRBPID) 1OPA 3	
247	1pmp	A	162	294	3.4e-53			76.11	CELLULAR LIPOPHILIC TRANSPORT PROTEIN P2 MYELIN PROTEIN (P2) 1PMP 3	
249	1tub	B	1	281	0			321.45	TUBULIN; CHAIN: A, B;	MICROTUBULES MICROTUBULES, ALPHA- TUBULIN, BETA-TUBULIN, GTPASE HELIX
255	1aoa		22	179	1.2e-20	0.41	0.95		T-FIMBRIN; CHAIN: NULL;	ACTIN-BINDING PROTEIN ACTIN-BINDING PROTEIN, CALCIUM-BINDING,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
255	1aoa		9	133	8.5e-06	-0.09	0.04		T-FIMBRIN; CHAIN: NULL;	PHOSPHORYLATION ACTIN-BINDING PROTEIN ACTIN-BINDING PROTEIN, CALCIUM-BINDING, PHOSPHORYLATION
255	1bkr	A	27	137	0.00068	0.03	0.07		SPECTRIN BETA CHAIN; CHAIN: A;	ACTIN-BINDING CALPONIN HOMOLOGY (CH) DOMAIN; FILAMENTOUS ACTIN- BINDING DOMAIN, CYTOSKELETON
255	1dxx	A	20	179	3.4e-31	0.21	-0.07		DYSTROPHIN; CHAIN: A, B, C, D;	STRUCTURAL PROTEIN DYSTROPHIN, MUSCULAR DYSTROPHY, CALPONIN HOMOLOGY DOMAIN, 2
255	1qag	A	24	179	6.8e-30	0.03	-0.05		UTROPHIN ACTIN BINDING REGION; CHAIN: A, B;	ACTIN-BINDING, UTROPHIN STRUCTURAL PROTEIN CALPONIN HOMOLOGY DOMAIN, DOMAIN SWAPPING, ACTIN BINDING, 2 UTROPHIN, DYSTROPHIN, STRUCTURAL PROTEIN
260	1mub	A	220	435	3.4e-20			60.57	BASEMENT MEMBRANE PROTEIN BM-40; CHAIN: A, B;	EXTRACELLULAR MODULE OSTEONECTIN, SPARC, SECRETED PROTEIN ACIDIC AND EXTRACELLULAR MODULE, GLYCOPROTEIN, 2 ANTI-ADHESIVE PROTEIN, 2 COLLAGEN BINDING, SITE- DIRECTED MUTAGENESIS, GLYCOSYLATED 3 PROTEIN MODRES
262	1bq0		275	340	5.1e-27	0.49	0.96		DNAI; CHAIN: NULL;	CHAPERONE HSP40; CHAPERONE, HEAT SHOCK, PROTEIN FOLDING, DNAK
262	1bdj		276	341	3.4e-27	0.65	1.00		HUMAN HSP40; CHAIN: NULL;	MOLECULAR CHAPERONE HDI-1; MOLECULAR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										CHAPERONE
263	1crz	A	18	202	1.7e-05	0.09	0.00		TOLB PROTEIN; CHAIN: A;	TOXIN BINDING PROTEIN TWO DOMAINS: BETA PROPELLER AND ALPHA/BETA FOLD
263	1erj	A	104	433	6.8e-63	0.03	0.23		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
263	1erj	A	10	316	1.4e-69	0.02	1.00		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
263	1erj	A	401	682	1.2e-61	0.04	-0.18		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
263	1erj	A	498	760	1.4e-56	-0.00	-0.18		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
263	1got	B	1	314	1e-78	0.38	0.76		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
263	1got	B	398	679	6.8e-69	0.17	-0.13		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
263	1got	B	498	761	6.8e-48	0.13	-0.19		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									GAMMA; CHAIN: G;	SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
263	1got	B	51	355	5.1e-64	0.28	0.19		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
263	1got	B	92	430	3.4e-57	0.10	-0.05		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
263	1qks	A	4	221	1.2e-22	0.13	-0.17		CYTOCHROME C1 NITRITE REDUCTASE; CHAIN: A, B;	OXIDOREDUCTASE ENZYME, NITRITE REDUCTASE, OXIDOREDUCTASE, DENITRIFICATION, 2 ELECTRON TRANSPORT, PERIPLASMIC
266	1fk	U	18	86	3.6e-34	-0.28	0.94		23S RRNA; CHAIN: 0; 5S RRNA; CHAIN: 9; RIBOSOMAL PROTEIN L2; CHAIN: A; RIBOSOMAL PROTEIN L3; CHAIN: B; RIBOSOMAL PROTEIN L4; CHAIN: C; RIBOSOMAL;	RIBOSOME 50S RIBOSOMAL PROTEIN L2P, HMAL2, HL4; 50S RIBOSOMAL PROTEIN L3P, HMAL3, HL1; 50S RIBOSOMAL PROTEIN L4E, HMAL4, HL6; 50S RIBOSOMAL PROTEIN L5P,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									PROTEIN L5; CHAIN: D; RIBOSOMAL PROTEIN L7AE; CHAIN: E; RIBOSOMAL PROTEIN L10E; CHAIN: F; RIBOSOMAL PROTEIN L13; CHAIN: G; RIBOSOMAL PROTEIN L14; CHAIN: H; RIBOSOMAL PROTEIN L15E; CHAIN: I; RIBOSOMAL PROTEIN L15; CHAIN: J; RIBOSOMAL PROTEIN L18; CHAIN: K; RIBOSOMAL PROTEIN L18E; CHAIN: L; RIBOSOMAL PROTEIN L19; CHAIN: M; RIBOSOMAL PROTEIN L21E; CHAIN: N; RIBOSOMAL PROTEIN L22; CHAIN: O; RIBOSOMAL PROTEIN L23; CHAIN: P; RIBOSOMAL PROTEIN L24; CHAIN: Q; L24E; CHAIN: R; RIBOSOMAL PROTEIN L29; CHAIN: S; RIBOSOMAL PROTEIN L30; CHAIN: T; RIBOSOMAL PROTEIN L31E; CHAIN: U; RIBOSOMAL PROTEIN L32E; CHAIN: V; RIBOSOMAL PROTEIN L37AE; CHAIN: W;	HMAL5, HL13; 30S RIBOSOMAL PROTEIN HS6; 50S RIBOSOMAL PROTEIN L13P, HMAL13; 50S RIBOSOMAL PROTEIN L14P, HMAL14, HL27; 50S RIBOSOMAL PROTEIN L15P, HMAL15, HL9; 50S RIBOSOMAL PROTEIN L18P, HMAL18, HL12; 50S RIBOSOMAL PROTEIN L18E, HL29, L19; 50S RIBOSOMAL PROTEIN L19E, HMAL19, HL24; 50S RIBOSOMAL PROTEIN L21E, HL31; 50S RIBOSOMAL PROTEIN L22P, HMAL22, HL23; 50S RIBOSOMAL PROTEIN L23P, HMAL23, HL25, L21; 50S RIBOSOMAL PROTEIN L24P, HMAL24, HL16, HL15; 50S RIBOSOMAL PROTEIN L24E, HL21/HL22; 50S RIBOSOMAL PROTEIN L29P, HMAL29, HL33; 50S RIBOSOMAL PROTEIN L30P, HMAL30, HL20, HL16; 50S RIBOSOMAL PROTEIN L31E, L34, HL30; 50S RIBOSOMAL PROTEIN L32E, HL5; 50S RIBOSOMAL PROTEIN L37E, L35E; 50S RIBOSOMAL PROTEINS L39E, HL39E, HL46E; 50S RIBOSOMAL PROTEIN L44E, LA, HLA; 50S RIBOSOMAL PROTEIN L6P, HMAL6, HL10 RIBOSOME ASSEMBLY, RNA-RNA, PROTEIN-RNA, PROTEIN-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									RIBOSOMAL PROTEIN L37E; CHAIN: X; RIBOSOMAL PROTEIN L39E; CHAIN: Y; RIBOSOMAL PROTEIN L44E; CHAIN: Z; RIBOSOMAL PROTEIN L6; CHAIN: 1	PROTEIN
266	1ffk	U	18	95	8.5e-21	-0.22	0.70		23S RRNA; CHAIN: 0; 5S RNA; CHAIN: 9; RIBOSOMAL PROTEIN L2; CHAIN: A; RIBOSOMAL PROTEIN L3; CHAIN: B; RIBOSOMAL PROTEIN L4; CHAIN: C; RIBOSOMAL PROTEIN L5; CHAIN: D; RIBOSOMAL PROTEIN L7AE; CHAIN: E; RIBOSOMAL PROTEIN L10E; CHAIN: F; RIBOSOMAL PROTEIN L13; CHAIN: G; RIBOSOMAL PROTEIN L14; CHAIN: H; RIBOSOMAL PROTEIN L15E; CHAIN: I; RIBOSOMAL PROTEIN L15; CHAIN: J; RIBOSOMAL PROTEIN L18; CHAIN: K; RIBOSOMAL PROTEIN L18E; CHAIN: L; L19; CHAIN: M; RIBOSOMAL PROTEIN L21E; CHAIN: N; RIBOSOMAL PROTEIN L22; CHAIN: O; RIBOSOMAL PROTEIN	RIBOSOME 50S RIBOSOMAL PROTEIN L2P, HMAL2, HL4; 50S RIBOSOMAL PROTEIN L3P, HMAL3, HL1; 50S RIBOSOMAL PROTEIN L4E, HMAL4, HL6; 50S RIBOSOMAL PROTEIN L5P, HMAL5, HL13; 30S RIBOSOMAL PROTEIN HS6; 50S RIBOSOMAL PROTEIN L13P, HMAL13; 50S RIBOSOMAL PROTEIN L14P, HMAL14, HL27; 50S RIBOSOMAL PROTEIN L15P, HMAL15, HL9; 50S RIBOSOMAL PROTEIN L18P, HMAL18, HL12; 50S RIBOSOMAL PROTEIN L18E, HL29, L19; 50S RIBOSOMAL PROTEIN L19E, HMAL19, HL24; 50S RIBOSOMAL PROTEIN L21E, HL31; 50S RIBOSOMAL PROTEIN L22P, HMAL22, HL23; 50S RIBOSOMAL PROTEIN L23P, HMAL23, HL25, L21; 50S RIBOSOMAL PROTEIN L24P, HMAL24, HL16, HL15; 50S RIBOSOMAL PROTEIN L24E, HL21/HL22; 50S RIBOSOMAL PROTEIN L29P, HMAL29.

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									L23; CHAIN: P; RIBOSOMAL PROTEIN L24; CHAIN: Q; RIBOSOMAL PROTEIN L24E; CHAIN: R; RIBOSOMAL PROTEIN L29; CHAIN: S; RIBOSOMAL PROTEIN L30; CHAIN: T; RIBOSOMAL PROTEIN L31E; CHAIN: U; RIBOSOMAL PROTEIN L32E; CHAIN: V; RIBOSOMAL PROTEIN L37AE; CHAIN: W; RIBOSOMAL PROTEIN L37E; CHAIN: X; RIBOSOMAL PROTEIN L39E; CHAIN: Y; RIBOSOMAL PROTEIN L44E; CHAIN: Z; RIBOSOMAL PROTEIN L6; CHAIN: 1;	HL33; 50S RIBOSOMAL PROTEIN L30P, HMAL30, HL20, HL16; 50S RIBOSOMAL PROTEIN L31E, L34, HL30; 50S RIBOSOMAL PROTEIN L32E, HL5; 50S RIBOSOMAL PROTEIN L37E, L35E; 50S RIBOSOMAL PROTEINS L39E, HL39E, HL46E; 50S RIBOSOMAL PROTEIN L44E, LA, HLA; 50S RIBOSOMAL PROTEIN L6P, HMAL6, HL10 RIBOSOME ASSEMBLY, RNA-RNA, PROTEIN-RNA, PROTEIN- PROTEIN
268	laye		119	548	1.7e-91			107.01	PROCARBOXYPEPTIDASE A2; CHAIN: NULL;	SERINE PROTEASE PCPA2; SERINE PROTEASE, ZYMOGEN, HYDROLASE
268	lnsa		123	549	1.7e-88			110.63	PROCARBOXYPEPTIDASE B; CHAIN: NULL	SERINE PROTEASE PORCINE PROCARBOXYPEPTIDASE, SERINE PROTEASE
268	lpca		122	554	3.4e-90			114.78	HYDROLASE(C- TERMINAL PEPTIDASE) PROCARBOXYPEPTIDASE A (E.C.3.4.12.2) IPCA 3	
269	lbw6	A	6	56	5.1e-07	0.08	0.83		CENTROMERE PROTEIN B; CHAIN: A;	DNA BINDING PROTEIN CENTROMERE PROTEIN, DNA-BINDING, HELIX-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										TURN-HELIX, DNA 2 BINDING PROTEIN
270	1alh	A	224	301	8.5e-21	-0.37	0.27		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
270	1alh	A	333	412	1.7e-28	-0.00	0.81		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
270	1alh	A	361	442	1.7e-28	-0.17	0.11		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
270	1alh	A	361	466	1.2e-17	-0.46	0.25		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
270	lard		252	278	7.2e-07	-0.13	0.00		TRANSCRIPTION REGULATION YEAST ADR1 (RESIDUES 102 - 130) 1ARD 3 (AMINO TERMINAL ZINC FINGER DOMAIN) (NMR, 10 STRUCTURES) 1ARD 4 (ADRI) 1ARD 5	
270	lard		252	280	6.8e-05	-0.24	0.37		TRANSCRIPTION REGULATION YEAST TRANSCRIPTION FACTOR	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									ADR1 (RESIDUES 102 - 130) 1ARD 3 (AMINO TERMINAL ZINC FINGER DOMAIN) (NMR, 10 STRUCTURES) 1ARD 4 (ADR1B) 1ARD 5	
270	1bbo		252	278	3.6e-07	-0.22	0.40		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11 IBBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) IBBO 4	
270	1bbo		362	408	3.6e-13	-0.46	0.58		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11 IBBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) IBBO 4	
270	1buo	A	1	117	1.7e-37	0.26	0.71		PROMYELOCYTIC LEUKEMIA ZINC FINGER PROTEIN PLZF; CHAIN: A;	GENE REGULATION POZ DOMAIN; PROTEIN-PROTEIN INTERACTION DOMAIN, TRANSCRIPTIONAL 2 REPRESSOR, ZINC-FINGER PROTEIN, X-RAY CRYSTALLOGRAPHY, 3 PROTEIN STRUCTURE, PROMYELOCYTIC LEUKEMIA, GENE REGULATION
270	1mey	C	223	301	3.4e-36	-0.31	0.33		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
270	1mey	C	332	412	1.7e-48	-0.13	0.49		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
270	1mey	C	388	470	8.5e-49	-0.06	0.16		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
270	1mey	G	249	276	1.7e-12	-0.45	0.57		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
270	1mey	G	358	385	3.4e-13	-0.14	0.39		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
270	1sp1		252	278	3.6e-05	0.03	0.43		SPIF3; CHAIN: NULL;	ZINC FINGER TRANSCRIPTION FACTOR SPI1; ZINC FINGER, TRANSCRIPTION ACTIVATION, SPI
270	1sp2		252	280	1.2e-07	-0.40	0.54		SPIF2; CHAIN: NULL;	ZINC FINGER TRANSCRIPTION FACTOR SPI1; ZINC FINGER, TRANSCRIPTION ACTIVATION, SPI
270	1tf3	A	333	412	5.1e-19	0.15	0.07		TRANSCRIPTION FACTOR	COMPLEX (TRANSCRIPTION)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
270	1tf3	A	361	466	2.4e-12	-0.60	0.25		IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
270	1tf6	A	333	479	3.4e-37	-0.50	0.03		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
270	1tf6	A	361	507	1.7e-36	-0.35	0.33		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
270	1ubd	C	340	442	3.4e-34	-0.28	0.29		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
270	2adr		231	278	3.4e-11	-0.69	0.12		ADR1; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADR1, ZINC FINGER, NMR
270	2gli	A	224	275	1.7e-07	-0.15	0.09		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
270	2gli	A	361	466	1.2e-15	-0.53	0.11		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
272	1b3u	A	478	888	1.2e-27	-0.30	0.21		PROTEIN PHOSPHATASE PP2A; CHAIN: A, B;	SCAFFOLD PROTEIN SCAFFOLD PROTEIN, PP2A, PHOSPHORYLATION, HEAT REPEAT
272	1ee4	A	436	838	1.4e-47	0.16	0.24		KARYOPHERIN ALPHA; CHAIN: A, B; MYC PROTO-ONCOGENE PROTEIN; CHAIN: C, D, E, F;	TRANSPORT PROTEIN SERINE-RICH RNA POLYMERASE I SUPPRESSOR PROTEIN; ARM REPEAT
272	1lal	A	437	838	8.5e-50	0.03	0.25		IMPORTIN ALPHA; CHAIN: A;	NUCLEAR IMPORT RECEPTOR KARYOPHERIN ALPHA; NUCLEAR IMPORT RECEPTOR, NUCLEAR LOCALIZATION SIGNAL, 2 ARMADILLO REPEATS, AUTOINHIBITION, INTRASTERIC REGULATION
272	1ibr	B	110	574	3.4e-20	0.06	0.07		RAN; CHAIN: A, C;;	SMALL GTPASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
272	1qbk	B	3	858	1.7e-57			116.75	IMPORTIN BETA SUBUNIT; CHAIN: B, D	KARYOPHERIN BETA, P95 SMALL GTPASE, NUCLEAR TRANSPORT RECEPTOR
272	1qgr	A	109	894	1.4e-39	-0.12	0.75		KARYOPHERIN BETA2; CHAIN: B; RAN; CHAIN: C;	NUCLEAR TRANSPORT PROTEIN COMPLEX HEAT REPEATS, NUCLEAR TRANSPORT PROTEIN COMPLEX
272	1qgr	A	10	849	1.4e-39			127.17	IMPORTIN BETA SUBUNIT; CHAIN: A; IMPORTIN ALPHA-2 SUBUNIT; CHAIN: B;	TRANSPORT RECEPTOR KARYOPHERIN BETA-1, NUCLEAR FACTOR P97, IMPORTIN IMPORTIN ALPHA-2 SUBUNIT, KARYOPHERIN ALPHA-2 TRANSPORT RECEPTOR, NUCLEAR IMPORT, HEAT MOTIF, NLS-BINDING
272	2bct		239	709	2.4e-17	-0.06	0.07		BETA-CATENIN; CHAIN: NULL;	STRUCTURAL PROTEIN ARMADILLO REPEAT, BETA-CATENIN, STRUCTURAL PROTEIN
273	1b3u	A	196	850	1.2e-72			125.57	PROTEIN PHOSPHATASE PP2A; CHAIN: A, B;	SCAFFOLD PROTEIN SCAFFOLD PROTEIN, PP2A, PHOSPHORYLATION, HEAT REPEAT
273	1b3u	A	207	847	1.2e-72	-0.02	0.69		PROTEIN PHOSPHATASE PP2A; CHAIN: A, B;	SCAFFOLD PROTEIN SCAFFOLD PROTEIN, PP2A, PHOSPHORYLATION, HEAT

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
273	1b3u	A	60	674	1.2e-46	0.02	0.92		PROTEIN PHOSPHATASE PP2A; CHAIN: A, B;	REPEAT SCAFFOLD PROTEIN SCAFFOLD PROTEIN, PP2A, PHOSPHORYLATION, HEAT REPEAT
273	1ee4	A	484	846	5.1e-51	-0.03	0.15		KARYOPHERIN ALPHA; CHAIN: A, B; MYC PROTO-ONCOGENE PROTEIN; CHAIN: C, D, E, F;	TRANSPORT PROTEIN SERINE-RICH RNA POLYMERASE I SUPPRESSOR PROTEIN; ARM REPEAT
273	1lal	A	483	846	5.1e-53	-0.20	0.66		IMPORTIN ALPHA; CHAIN: A;	NUCLEAR IMPORT RECEPTOR KARYOPHERIN ALPHA; NUCLEAR IMPORT RECEPTOR, NUCLEAR LOCALIZATION SIGNAL, 2 ARMADILLO REPEATS, AUTOINHIBITION, INTRASTERIC REGULATION
273	1qbk	B	3	824	1e-43			125.55	KARYOPHERIN BETA2; CHAIN: B; RAN; CHAIN: C;	NUCLEAR TRANSPORT PROTEIN COMPLEX HEAT REPEATS, NUCLEAR TRANSPORT PROTEIN COMPLEX
273	1qgr	A	111	856	3.4e-37	-0.31	0.05		IMPORTIN BETA SUBUNIT; CHAIN: A; IMPORTIN ALPHA-2 SUBUNIT; CHAIN: B;	TRANSPORT RECEPTOR KARYOPHERIN BETA-1, NUCLEAR FACTOR P97, IMPORTIN IMPORTIN ALPHA-2 SUBUNIT, KARYOPHERIN ALPHA-2
									TRANSPORT RECEPTOR, NUCLEAR IMPORT, HEAT MOTIF, NLS-BINDING	
274	1al7		259	362	0.0012	-0.59	0.00		SERINE/THREONINE PROTEIN PHOSPHATASE 5; CHAIN: NULL;	HYDROLASE TETRATRICOPEPTIDE, TRP- HYDROLASE, PHOSPHATASE, PROTEIN- PROTEIN INTERACTIONS, TPR, 2 SUPER-HELIX, X-RAY

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										STRUCTURE
275	1bz4	A	131	251	0.00017	-0.14	0.27		APOLIPOPROTEIN E; CHAIN: A;	LIPID BINDING PROTEIN APO-E3; LIPID TRANSPORT, LIPID TRANSPORT, HEPARIN-BINDING, PLASMA 2 PROTEIN, HDL, VLDL REMARK
275	1dkg	A	158	289	0.00012	-0.54	0.18		NUCLEOTIDE EXCHANGE FACTOR GRPE; CHAIN: A, B; MOLECULAR CHAPERONE DNAK; CHAIN: D;	COMPLEX (HSP24/HSP70) HSP70, GRPE, MOLECULAR CHAPERONE, NUCLEOTIDE EXCHANGE 2 FACTOR, COILED-COIL, COMPLEX (HSP24/HSP70)
275	1dn1	B	143	268	2.4e-06	-0.00	0.00		SYNTAXIN BINDING PROTEIN 1; CHAIN: A; SYNTAXIN 1A; CHAIN: B; SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS S NSEC1; PROTEIN-PROTEIN COMPLEX, MULTI-SUBUNIT
275	1ez3	A	145	270	1.2e-07	0.21	0.11			ENDOCYTOSIS/EXOCYTOSIS S SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
275	1fre		106	143	9.6e-14	-0.40	0.54		NUCLEAR FACTOR XNF7; CHAIN: NULL;	ZINC-BINDING PROTEIN, ZINC-BINDING PROTEIN, XNF7, BBOX, DEVELOPMENT, 3 MID-BLASTULA-TRANSITION
275	1qum	A	148	268	1.2e-07	0.16	0.51		HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
275	1qum	A	85	349	1.2e-07			59.40	HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
276	1av1	A	224	427	2.4e-10			64.70	APOLIPOPROTEIN A-I; CHAIN: A, B, C, D;	LIPID TRANSPORT APO A-I; LIPOPROTEIN, LIPID TRANSPORT,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										CHOLESTEROL METABOLISM, 2 ATHEROSCLEROSIS, HDL, LCAT-ACTIVATION
276	1ez3	A	228	332	4.8e-08	0.19	-0.19		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
276	1ez3	A	239	358	1.2e-10	0.09	0.01		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
276	1quu	A	180	445	1.2e-24			65.07	HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
276	1sig		166	469	4.8e-09			77.03	RNA POLYMERASE PRIMARY SIGMA FACTOR; CHAIN: NULL;	TRANSCRIPTION REGULATION SIGMA70; RNA POLYMERASE SIGMA FACTOR, TRANSCRIPTION REGULATION
276	1sig		291	498	1.2e-06	-0.16	0.06		RNA POLYMERASE PRIMARY SIGMA FACTOR; CHAIN: NULL;	TRANSCRIPTION REGULATION SIGMA70; RNA POLYMERASE SIGMA FACTOR, TRANSCRIPTION REGULATION
277	1buo	A	26	149	3.4e-21	0.56	1.00		PROMYELOCYTIC LEUKEMIA ZINC FINGER PROTEIN PLZF; CHAIN: A;	GENE REGULATION POZ DOMAIN; PROTEIN-PROTEIN INTERACTION DOMAIN, TRANSCRIPTIONAL 2 REPRESSOR, ZINC-FINGER PROTEIN, X-RAY CRYSTALLOGRAPHY, 3 PROTEIN STRUCTURE, PROMYELOCYTIC

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
277	1buo	A	31	146	2.4e-22	0.28	1.00		PROMYELOCYTIC LEUKEMIA ZINC FINGER PROTEIN PLZF; CHAIN: A;	LEUKEMIA, GENE REGULATION
277	1gof		299	611	8.5e-13	0.19	0.19		OXIDOREDUCTASE(OXYGEN(A)) GALACTOSE OXIDASE (E.C.1.1.3.9) (PH 4.5) 1GOF 3	GENE REGULATION POZ DOMAIN; PROTEIN-PROTEIN INTERACTION DOMAIN, TRANSCRIPTIONAL 2 REPRESSOR, ZINC-FINGER PROTEIN, X-RAY CRYSTALLOGRAPHY, 3 PROTEIN STRUCTURE, PROMYELOCYTIC LEUKEMIA, GENE REGULATION
278	1alh	A	120	215	5.1e-22	0.24	0.87		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
278	1alh	A	148	243	3.4e-23	0.21	0.13		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
278	1alh	A	152	244	6e-29	-0.40	0.09		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
278	1mey	C	119	215	1e-38	-0.02	0.10		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER	COMPLEX (ZINC FINGER/DNA) ZINC FINGER

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
278	1mey	C	147	243	6.8e-41	0.19	0.89		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
278	1mey	C	190	271	1e-44	0.29	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
278	1mey	C	218	299	3.4e-46	0.30	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
278	1mey	C	246	327	6.8e-47	0.21	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
278	1mey	C	274	355	1.7e-48	0.37	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
278	1mey	C	302	383	6.8e-49	0.48	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	(ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
278	1mey	C	330	411	1.7e-49	0.55	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
278	1mey	C	358	439	3.4e-50	0.31	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
278	1mey	C	386	467	8.5e-51	0.32	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
278	1mey	C	414	495	1.4e-50	0.52	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
278	1mey	C	442	523	1.7e-50	0.38	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
278	1mey	C	442	524	1.4e-50			100.48	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
278	1mey	C	470	551	3.4e-50	0.43	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
278	1mey	C	498	579	8.5e-50	0.61	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
278	1mey	C	526	585	1.7e-35	0.48	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
278	1mey	C	97	172	5.1e-37	0.11	0.09		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
278	1mcy	G	119	144	1.2e-07	-0.05	0.07		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
278	1f3	A	120	211	8.5e-16	-0.05	0.00		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
278	1f3	A	148	243	3.4e-15	0.08	-0.02		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
278	1df6	A	148	308	1.2e-32	0.05	0.64		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
278	1df6	A	219	364	1.7e-36	0.10	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
278	1tf6	A	275	420	5.1e-37	-0.03	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
278	1tf6	A	386	554	1.7e-37			104.66	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
278	1tf6	A	387	539	1.7e-37	0.35	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
278	1tf6	A	415	561	3.4e-37	0.18	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
278	1tf6	A	443	581	1.4e-35	0.22	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
278	1ubd	C	105	215	3.4e-25	-0.01	0.23		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
278	1ubd	C	155	271	3.4e-28	-0.03	0.93		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
278	1ubd	C	198	299	8.5e-32	-0.27	0.92		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
278	1ubd	C	282	383	6.8e-34	0.10	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Pst Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
278	1ubd	C	310	411	1.5e-34	0.30	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
278	1ubd	C	366	467	1e-34	0.32	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
278	1ubd	C	416	524	3.4e-35			83.90	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
278	1ubd	C	422	523	3.4e-35	0.14	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
278	2gli	A	129	245	3.6e-18	-0.42	0.03		ZINC FINGER PROTEIN	COMPLEX (DNA-BINDING)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Pst Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
278	2gli	A	144	329	2.4e-52	0.23	0.72		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
278	2gli	A	190	326	3.4e-32	0.40	0.99		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
278	2gli	A	246	413	6e-68	0.07	0.99		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
278	2gli	A	274	413	6e-68			94.24	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
278	2gli	A	282	410	1.7e-34	0.23	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
278	2gli	A	366	494	3.4e-34	0.16	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
278	2gli	A	386	553	3.6e-71	0.14	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
278	2gli	A	442	581	2.4e-73	0.29	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
279	1alh	A	137	207	5.1e-11	0.15	-0.20		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
279	1alh	A	161	239	1.5e-15	0.05	-0.20		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
279	1alh	A	299	397	1.5e-23	0.21	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
279	1alh	A	457	538	1.4e-22			86.47	QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
279	1alh	A	484	542	1.4e-22	0.39	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
279	1bbo		431	484	8.4e-24	-0.10	0.36		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
279	1mey	C	160	239	1.7e-30	0.01	-0.20		MUTANT WITH CYS 11 IBBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) IBBO 4	
279	1mey	C	298	397	1.5e-42	0.14	0.95		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
279	1mey	C	345	425	3.4e-47	0.48	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
279	1mey	C	372	453	1e-49	0.57	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
279	1mey	C	400	480	3.4e-48	0.79	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
279	1mey	C	456	536	1.7e-48	0.51	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
279	1mey	C	456	537	1.7e-48			107.34	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
279	1mey	C	483	542	6.8e-36	0.61	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
279	1ef3	A	146	200	1.7e-08	0.22	-0.20		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
279	1ef3	A	161	238	5.1e-11	0.03	-0.20		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
279	1tf3	A	346	456	1.1e-24	0.05	0.03		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	(TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
279	1tf6	A	346	496	6.8e-35	0.24	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
279	1tf6	A	372	537	8.5e-37			115.39	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
279	1tf6	A	401	538	8.5e-37	0.28	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
279	1ubd	C	328	425	3.4e-29	0.22	1.00		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									DNA; CHAIN: A, B;	INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
279	1ubd	C	353	453	1.2e-34	0.05	1.00		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
279	1ubd	C	408	508	3.4e-34	0.55	1.00		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
279	1ubd	C	428	537	3.4e-34			86.73	YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
279	1ubd	C	436	536	8.5e-33	0.23	1.00		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
279	2adr		271	343	1e-11	0.01	-0.18		ADR1; CHAIN: NULL;	PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
279	2gli	A	187	343	8.5e-25	0.06	-0.18		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADR1, ZINC FINGER, NMR COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
279	2gli	A	333	452	1.7e-30	0.28	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
279	2gli	A	346	510	7.2e-63	0.04	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
279	2gli	A	372	510	7.2e-63			106.71	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
279	2gli	A	380	507	6.8e-34	0.57	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
279	2gli	A	401	537	1.2e-61	0.53	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
279	2gli	A	408	97	1.2e-33	0.57	1.00		ZINC FINGER PROTEIN	COMPLEX (DNA-BINDING PROTEIN/DNA)

Table 5

SEQ ID NO;	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
279	2gli	A	436	543	3.4e-30	0.46	0.92		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING; PROTEIN/DNA)
282	1f5a	A	19	421	5.1e-95	0.78	1.00		POLY(A) POLYMERASE; CHAIN: A;	TRANSFERASE MRNA PROCESSING, TRANSFERASE, TRANSCRIPTION, RNA-BINDING, 2 PHOSPHORYLATION, NUCLEAR PROTEIN, ALTERNATIVE SPLICING 3 HELICAL TURN MOTIF, NUCLEOTIDYL TRANSFERASE CATALYTIC DOMAIN
286	1mey	C	367	449	1.4e-50			103.07	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
286	1tf6	A	283	446	6.8e-38			110.63	TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
286	1ubd	C	116	225	5.1e-33			87.81	YY1; CHAIN: C; ADENO-	COMPLEX (TRANSCRIPTION)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
286	2gli	A	255	394	6.8e-35			96.09	ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	REGULATION/DNA YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
287	1a1h	A	138	233	3.4e-21	-0.13	0.33		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
287	1a1h	A	146	234	3.6e-25	0.29	0.43		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
287	1a1h	A	489	597	1.2e-25	-0.32	0.12		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
287	1mev	C	116	205	6.8e-30	0.09	0.60		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2-CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
287	1mey	C	137	233	5.1e-39	0.23	0.89		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
287	1mey	C	180	261	1.7e-44	0.23	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
287	1mey	C	208	289	1.7e-46	0.24	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
287	1mey	C	208	290	6.8e-48			103.66	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
287	1mey	C	236	317	6.8e-48	0.10	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
287	1mey	C	264	345	3.4e-49	0.48	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
287	1mey	C	292	373	1.4e-49	0.11	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
287	1mey	C	320	401	8.5e-50	0.25	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
287	1mey	C	348	429	1.7e-50	0.12	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
287	1mey	C	376	457	5.1e-51	0.03	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
287	1mey	C	404	485	3.4e-51	0.25	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
287	1mey	C	432	513	6.8e-51	0.24	1.00		DNA; CHAIN: A, B, D, E;	COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
287	1mey	C	460	569	5.1e-46	-0.25	0.99		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
287	1mey	C	516	597	5.1e-39	0.26	0.93		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
287	1mey	C	544	625	6.8e-50	0.15	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
287	1mey	C	572	653	1e-49	0.32	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
287	1mey	G	318	345	5.1e-13	0.56	0.95		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
287	1sp2		489	517	1.7e-08	-0.38	0.66		SPIP2; CHAIN: NULL;	STRUCTURE, COMPLEX (ZINC FINGER/DNA) ZINC FINGER TRANSCRIPTION FACTOR SPI; ZINC FINGER, TRANSCRIPTION ACTIVATION, SPI
287	1tf3	A	489	597	5.1e-16	-0.50	0.07		TRANSCRIPTION FACTOR IIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
287	1tf6	A	138	298	1e-30	0.07	0.94		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE II, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
287	1tf6	A	181	326	5.1e-35	-0.10	0.93		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE II, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
287	1tf6	A	208	382	3.4e-37			108.67	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE II, 2 TRANSCRIPTION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
287	1tf6	A	237	382	3.4e-37	-0.04	0.96		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	INITIATION, ZINC FINGER PROTEIN COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
287	1tf6	A	293	443	8.5e-37	-0.10	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
287	1tf6	A	377	522	3.4e-38	0.03	0.90		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
287	1tf6	A	405	578	1e-33	-0.35	0.57		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
287	1tf6	A	489	653	1.2e-31	-0.22	0.43		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
287	1ubd	C	124	233	1.2e-21	-0.42	0.30		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
287	1ubd	C	140	261	1.5e-27	0.01	1.00		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
287	1ubd	C	183	289	1.7e-32	0.01	1.00		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
287	1ubd	C	244	345	1.7e-33	0.22	1.00		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
287	1ubd	C	272	373	5.1e-35	-0.10	1.00		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
287	1ubd	C	300	401	1.4e-34	-0.00	0.99		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
287	1ubd	C	356	457	8.5e-35	-0.23	0.98		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
287	1ubd	C	376	486	5.1e-36			92.78	YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
287	1ubd	C	384	485	5.1e-36	-0.09	0.95		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
287	1ubd	C	412	513	1.7e-34	-0.33	0.88		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
287	1ubd	C	468	569	1.2e-25	-0.04	0.84		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
287	1ubd	C	552	653	1.2e-33	0.02	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
287	2gli	A	137	260	1.7e-26	0.23	0.80		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
287	2gli	A	180	319	7.2e-67			92.67	ZINC FINGER PROTEIN	COMPLEX (DNA-BINDING PROTEIN/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
287	2gli	A	183	319	4.8e-63	0.28	0.95		GLI1; CHAIN: A; DNA; CHAIN: C, D;	PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
287	2gli	A	208	347	7.2e-67	0.08	0.94		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
287	2gli	A	264	459	3.6e-66	-0.03	0.33		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
287	2gli	A	272	400	6.8e-33	0.03	0.88		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
287	2gli	A	356	484	6.8e-34	0.04	0.81		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
287	2gli	A	404	627	3.6e-65	-0.22	0.11		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
287	2gli	A	488	654	2.4e-66	0.18	0.95		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
287	2gli	A	496	652	3.4e-32	0.11	0.89		ZINC FINGER PROTEIN GLI; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
288	1b6d	A	34	103	5.1e-34	-0.33	0.43		IMMUNOGLOBULIN; CHAIN: A, B;	IMMUNOGLOBULIN LIGHT-CHAIN DIMER HEADER
288	1bj1	L	34	103	3.4e-35	-0.05	0.17		FAB FRAGMENT; CHAIN: L, H, J, K; VASCULAR ENDOTHELIAL GROWTH FACTOR; CHAIN: V, W;	COMPLEX (ANTIBODY/ANTIGEN) FAB-12; VEGF; COMPLEX (ANTIBODY/ANTIGEN), ANGIOGENIC FACTOR
288	1bvk	A	34	103	3.4e-34	-0.56	0.47		HULYS11; CHAIN: A, B, D, E; LYSOZYME; CHAIN: C, F;	COMPLEX (HUMANIZED ANTIBODY/HYDROLASE) MURAMIDASE; HUMANIZED ANTIBODY, ANTIBODY COMPLEX, FV, ANTI-LYSOZYME, 2 COMPLEX (HUMANIZED ANTIBODY/HYDROLASE)
288	1bww	A	34	103	6.8e-35	-0.42	0.16		IG KAPPA CHAIN V-1 REGION REI; CHAIN: A, B;	IMMUNE SYSTEM REIV, STABILIZED IMMUNOGLOBULIN FRAGMENT, BENCE-JONES 2 PROTEIN, IMMUNE SYSTEM
288	1ce1	L	34	103	1.2e-33	-0.00	0.29		CAMPATH-1H LIGHT CHAIN; CHAIN: L; CAMPATH-1H HEAVY CHAIN; CHAIN: H; PEPTIDE ANTIGEN; CHAIN: P;	ANTIBODY THERAPEUTIC, ANTIBODY, CD52
288	1dee	A	34	103	3.4e-35	-0.12	0.27		IGM RF 2A2; CHAIN: A, C, E; IGM RF 2A2; CHAIN: B, D, F; IMMUNOGLOBULIN G BINDING PROTEIN A;	IMMUNE SYSTEM FAB-IBP COMPLEX CRYSTAL STRUCTURE 2.7A RESOLUTION BINDING 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
288	1dql	L	34	103	6.8e-34	-0.37	0.16		CHAIN: G, H; IGM MEZ IMMUNOGLOBULIN; CHAIN: L; IGM MEZ IMMUNOGLOBULIN; CHAIN: H; IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 IFGV 3 ANTIBODY 'H52' (HUH52-AA FV) IFGV 4 IMMUNOGLOBULIN IMMUNOGLOBULIN M (IG-M) FV FRAGMENT IGM 3	OUTSIDE THE ANTIGEN COMBINING SITE SUPERANTIGEN FAB VH3 3 SPECIFICITY IMMUNE SYSTEM IMMUNOGLOBULIN FOLD, ANTIBODY, IGM, FV
288	1fgv	L	34	103	1.7e-36	-0.36	0.12		IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 IFGV 3 ANTIBODY 'H52' (HUH52-AA FV) IFGV 4 IMMUNOGLOBULIN IMMUNOGLOBULIN M (IG-M) FV FRAGMENT IGM 3	
288	1igm	L	34	103	5.1e-34	-0.26	0.55		IMMUNOGLOBULIN IMMUNOGLOBULIN M (IG-M) FV FRAGMENT IGM 3	
288	2fgw	L	34	103	1.4e-36	-0.15	0.43		IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 2FGW 3 ANTIBODY 'H52' (HUH52-OZ FAB) 2FGW 4	
289	1drf		2	187	1.5e-77	0.81	1.00		DIHYDROFOLATE REDUCTASE (E.C.1.5.1.3) COMPLEX WITH FOLATE 1DRF 3	
289	1drf		3	187	1.5e-77			281.49	DIHYDROFOLATE REDUCTASE (E.C.1.5.1.3) COMPLEX WITH FOLATE 1DRF 3	
290	1dby	A	102	209	1.5e-22	0.19	0.09		CHLOROPLAST THIOREDOXIN M CH2; CHAIN: A;	OXIDOREDUCTASE THIOREDOXIN M, THIOREDOXIN CH2, CHLOROPLASTIC

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
290	1erv		103	209	3.4e-22	0.07	0.06		THIOREDOXIN; CHAIN: NULL;	THIOREDOXIN OXIDOREDUCTASE DIMER, THIOREDOXIN, X-RAY CRYSTALLOGRAPHY, OXIDOREDUCTASE
290	1f9m	A	97	207	1.5e-20	0.28	-0.03		THIOREDOXIN F; CHAIN: A, B;	ELECTRON TRANSPORT ELECTRON TRANSPORT
290	1faa	A	97	207	1.5e-20	0.10	0.04		THIOREDOXIN F; CHAIN: A;	ELECTRON TRANSPORT
290	1fb6	A	99	210	3.4e-23	0.46	-0.06		THIOREDOXIN M; CHAIN: A, B;	ELECTRON TRANSPORT ELECTRON TRANSPORT
290	1fp3	A	528	648	2.4e-06	0.25	0.42		N-ACYL-D- GLUCOSAMINE 2- EPIMERASE; CHAIN: A, B;	ISOMERASE ALPHA/ALPHA- BARREL, N-ACYL-D- GLUCOSAMINE 2- EPIMERASE
290	1quw	A	102	206	1.2e-26	0.35	0.06		THIOREDOXIN; CHAIN: A;	ELECTRON TRANSPORT ALPHA/BETA OPEN- TWISTED PROTEIN, THIOL- DISULFIDE
290	1r7p	B	99	209	6.8e-23	0.73	0.04		DNA POLYMERASE; CHAIN: A; THIOREDOXIN; CHAIN: B; DNA; CHAIN: P, T;	T7 DNA POLYMERASE, DNA REPLICATION, NUCLEOTIDYL 2 TRANSFERASE, SEQUENCING, THIOREDOXIN, PROCESSIVITY FACTOR, 3 COMPLEX (HYDROLASE/ELECTRON TRANSPORT/DNA)
290	1tof		98	210	3.4e-20	0.08	-0.01		THIOREDOXIN H; CHAIN: NULL;	ELECTRON TRANSPORT HTRX, HCH1, CH1; OXIDOREDUCTASE, ELECTRON TRANSPORT
290	2trx	A	97	209	3.4e-23	0.33	-0.07		ELECTRON TRANSPORT THIOREDOXIN 2TRXA 2 2TRXA 3	
291	1be9	A	36	83	3.6e-07	-0.13	0.37		PSD-95; CHAIN: A; CRPT; CHAIN: B;	PEPTIDE RECOGNITION PEPTIDE RECOGNITION,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
291	1pdr		10	90	2.4e-14	0.76	0.99		HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	PROTEIN LOCALIZATION SIGNAL TRANSDUCTION HDLG, DHR3 DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, REPEAT
291	1qau	A	13	102	3.6e-17	0.42	0.45		NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: A;	OXIDOREDUCTASE BETA-FINGER
291	1qav	A	13	83	9.6e-17	1.04	1.00		ALPHA-1 SYNTROPHIN (RESIDUES 77-171); CHAIN: A; NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: B;	MEMBRANE PROTEIN/OXIDOREDUCTASE BETA-FINGER, HETERODIMER
291	1qlc	A	10	87	7.2e-16	0.55	0.41		POSTSYNAPTIC DENSITY PROTEIN 95; CHAIN: A;	PEPTIDE RECOGNITION PSD-95; PDZ DOMAIN, NEURONAL NITRIC OXIDE SYNTHASE, NMDA RECEPTOR 2 BINDING
294	1a7a	A	34	79	0.001	-0.13	0.01		S-ADENOSYLHOMOCYSTEINE HYDROLASE; CHAIN: A, B;	HYDROLASE HYDROLASE, NAD BINDING PROTEIN
294	1ael	A	32	286	8.5e-63	0.10	1.00		TROPINONE REDUCTASE-I; CHAIN: A, B;	OXIDOREDUCTASE OXIDOREDUCTASE, TROPANE ALKALOID BIOSYNTHESIS, REDUCTION OF 2 TROPINONE TO TROPINE, SHORT-CHAIN DEHYDROGENASE
294	1ael	A	32	289	8.5e-63			84.17	TROPINONE REDUCTASE-I; CHAIN: A, B;	OXIDOREDUCTASE OXIDOREDUCTASE, TROPANE ALKALOID BIOSYNTHESIS, REDUCTION OF 2 TROPINONE TO TROPINE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
294	1ael	B	32	286	1.7e-65	0.16	1.00		TROPINONE REDUCTASE-1; CHAIN: A, B;	SHORT-CHAIN DEHYDROGENASE OXIDOREDUCTASE OXIDOREDUCTASE, TROPANE ALKALOID BIOSYNTHESIS, REDUCTION OF 2 TROPINONE TO TROPINE, SHORT-CHAIN DEHYDROGENASE
294	1ael	B	32	289	1.7e-65			89.11	TROPINONE REDUCTASE-1; CHAIN: A, B;	OXIDOREDUCTASE OXIDOREDUCTASE, TROPANE ALKALOID BIOSYNTHESIS, REDUCTION OF 2 TROPINONE TO TROPINE, SHORT-CHAIN DEHYDROGENASE
294	1bdb		33	308	1.7e-48			78.30	CIS-BIPHENYL-2,3-DIHYDRODIOL-2,3-DEHYDROGENASE; CHAIN: NULL;	OXIDOREDUCTASE NAD-DEPENDENT OXIDOREDUCTASE, SHORT-CHAIN ALCOHOL 2 DEHYDROGENASE, PCB DEGRADATION
294	1bdb		35	289	1.7e-48	0.43	1.00		CIS-BIPHENYL-2,3-DIHYDRODIOL-2,3-DEHYDROGENASE; CHAIN: NULL;	OXIDOREDUCTASE NAD-DEPENDENT OXIDOREDUCTASE, SHORT-CHAIN ALCOHOL 2 DEHYDROGENASE, PCB DEGRADATION
294	1cyd	A	33	287	1.7e-62	0.17	1.00		CARBONYL REDUCTASE; CHAIN: A, B, C, D;	OXIDOREDUCTASE SHORT-CHAIN DEHYDROGENASE, OXIDOREDUCTASE
294	1cyd	A	33	287	1.7e-62			84.06	CARBONYL REDUCTASE; CHAIN: A, B, C, D;	OXIDOREDUCTASE SHORT-CHAIN DEHYDROGENASE, OXIDOREDUCTASE
294	1eny		32	293	6.8e-19			51.90	ENOYL-ACYL CARRIER PROTEIN (ACP) REDUCTASE; 1ENY 4	OXIDOREDUCTASE INHA; 1ENY 6

Table 5

SEQ ID NO;	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
294	1fds		36	296	1.7e-47			112.71	CHAIN: NULL; 1ENY 5 17-BETA-HYDROXYSTEROID-DEHYDROGENASE; CHAIN: NULL;	DEHYDROGENASE DEHYDROGENASE, 17-BETA-HYDROXYSTEROID
294	1fds		39	269	1.7e-47	0.31	1.00		17-BETA-HYDROXYSTEROID-DEHYDROGENASE; CHAIN: NULL;	DEHYDROGENASE DEHYDROGENASE, 17-BETA-HYDROXYSTEROID
294	1fmc	A	29	306	1e-64			84.13	7 ALPHA-HYDROXYSTEROID DEHYDROGENASE; CHAIN: A, B;	OXIDOREDUCTASE SHORT-CHAIN DEHYDROGENASE/REDUCTASE, BILE ACID CATABOLISM
294	1fmc	A	33	264	1e-64	0.28	1.00		7 ALPHA-HYDROXYSTEROID DEHYDROGENASE; CHAIN: A, B;	OXIDOREDUCTASE SHORT-CHAIN DEHYDROGENASE/REDUCTASE, BILE ACID CATABOLISM
294	1hdc	A	33	284	1.7e-69	0.31	0.94		OXIDOREDUCTASE 3-ALPHA, 20-BETA-HYDROXYSTEROID DEHYDROGENASE (E.C.1.1.1.53) 1HDC 3 COMPLEXED WITH CARBENOXOLONE 1HDC 4	
294	1hdc	A	33	296	1.7e-69			84.15	OXIDOREDUCTASE 3-ALPHA, 20-BETA-HYDROXYSTEROID DEHYDROGENASE (E.C.1.1.1.53) 1HDC 3 COMPLEXED WITH CARBENOXOLONE 1HDC 4	
294	1osa		32	284	1.2e-32			58.34	SEPIAPTERIN REDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE SEPIAPTERIN REDUCTASE TETRAHYDROBIPTERIN OXIDOREDUCTASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
294	1oaa		36	225	1.2e-32	0.27	0.94		SEPIAPTERIN REDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE SEPIAPTERIN REDUCTASE, TETRAHYDROBIOPTERIN, OXIDOREDUCTASE
294	1ybv	A	25	284	1.7e-60			78.32	TRIHYDROXYNAPHTHALENE REDUCTASE; CHAIN: A, B;	OXIDOREDUCTASE NAPHTHOL REDUCTASE; OXIDOREDUCTASE
294	1ybv	A	34	267	1.7e-60	0.58	1.00		TRIHYDROXYNAPHTHALENE REDUCTASE; CHAIN: A, B;	OXIDOREDUCTASE NAPHTHOL REDUCTASE; OXIDOREDUCTASE
294	2ae2	A	31	301	3.4e-61			95.13	TROPINONE REDUCTASE-II; CHAIN: A, B;	OXIDOREDUCTASE OXIDOREDUCTASE, TROPANE ALKALOID BIOSYNTHESIS, REDUCTION OF 2 TROPINONE TO PSEUDOTROPINE, SHORT-CHAIN DEHYDROGENASE
294	2ae2	A	32	286	3.4e-61	0.25	1.00		TROPINONE REDUCTASE-II; CHAIN: A, B;	OXIDOREDUCTASE OXIDOREDUCTASE, TROPANE ALKALOID BIOSYNTHESIS, REDUCTION OF 2 TROPINONE TO PSEUDOTROPINE, SHORT-CHAIN DEHYDROGENASE
294	3grs		41	67	0.0085	-0.75	0.07		OXIDOREDUCTASE (FLAVOENZYMES) GLUTATHIONE REDUCTASE (E.C.1.6.4.2), OXIDIZED FORM (B) 3GRS	
295	lido		2	136	1.2e-09	0.16	0.95		INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
295	1lfa	A	2	142	1.2e-06	0.10	0.07		CD11A; ILFA 5 CHAIN: A, B; ILFA 6	CELL ADHESION LFA-1, ALPHA-1, BETA-2 INTEGRIN, A-DOMAIN; ILFA 8
298	1bg1	A	173	339	3.4e-05	0.02	0.01		STAT3B; CHAIN: A; 18-MER DESOXYOLIGONUCLEOTIDE; CHAIN: B;	COMPLEX (TRANSCRIPTION FACTOR/DNA) TRANSCRIPTION FACTOR, PROTEIN-DNA COMPLEX, CYTOKINE 2 ACTIVATION, COMPLEX (TRANSCRIPTION FACTOR/DNA)
300	1aab		239	294	4.8e-14	0.32	1.00		HIGH MOBILITY GROUP PROTEIN; 1AAB 5 CHAIN: NULL; 1AAB 6	DNA-BINDING HMGA DNA-BINDING HMG-BOX DOMAIN A OF RAT HMGI; 1AAB 8 HMG-BOX 1AAB 20
300	1aab		249	298	8.5e-05	0.67	0.99		HIGH MOBILITY GROUP PROTEIN; 1AAB 5 CHAIN: NULL; 1AAB 6	DNA-BINDING HMGA DNA-BINDING HMG-BOX DOMAIN A OF RAT HMGI; 1AAB 8 HMG-BOX 1AAB 20
300	1cg7	A	224	317	6.8e-15			54.43	NON HISTONE PROTEIN 6 A; CHAIN: A;	DNA BINDING PROTEIN HMG BOX, DNA BENDING, DNA RECOGNITION, CHROMATIN, NMR, DNA 2 BINDING PROTEIN
300	1cg7	A	249	299	6.8e-15	0.52	1.00		NON HISTONE PROTEIN 6 A; CHAIN: A;	DNA BINDING PROTEIN HMG BOX, DNA BENDING, DNA RECOGNITION, CHROMATIN, NMR, DNA 2 BINDING PROTEIN
300	1ckt	A	249	298	0.00014	0.63	0.99		HIGH MOBILITY GROUP 1 PROTEIN; CHAIN: A; DNA (5'-D(*CP*CP*(DDO) CHAIN: B; DNA (5'- CHAIN: C;	GENE REGULATION/DNA HMGI-1, AMPHOTERIN, HEPARIN-BINDING PROTEIN P30; HIGH-MOBILITY GROUP DOMAIN-BENT DNA, PROTEIN-DRUG DNA 2 COMPLEX, GENE REGULATION/DNA

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
300	1hme		240	292	2.4e-15	0.34	0.88		DNA-BINDING HIGH MOBILITY GROUP PROTEIN FRAGMENT-B (HMG1) (DNA-BINDING I HME 3 HMG-BOX DOMAIN B OF RAT HMG1) (NMR, 1 STRUCTURE) I HME 4	
300	1hme		250	298	1e-09	0.36	0.90		DNA-BINDING HIGH MOBILITY GROUP PROTEIN FRAGMENT-B (HMG1) (DNA-BINDING I HME 3 HMG-BOX DOMAIN B OF RAT HMG1) (NMR, 1 STRUCTURE) I HME 4	
300	1hry	A	249	299	1.7e-05	-0.08	0.16		HUMAN SRY; I HRY 6 CHAIN: A; I HRY 7 DNA; I HRY 9 CHAIN: B; I HRY 10	COMPLEX (DNA-BINDING PROTEIN/DNA)
300	1hsm		250	319	5.1e-10	0.10	0.41		DNA-BINDING HIGH MOBILITY GROUP PROTEIN 1 (HMG1) BOX 2, COMPLEXED WITH I HSM 3 MERCAPTOETHANOL (NMR, MINIMIZED AVERAGE STRUCTURE) I HSM 4	
300	1qrv	A	249	298	1.7e-05	0.51	0.87		DNA (5'-D(*GP*CP*GP*AP*TP*AP*TP*CP*GP*C)-3); CHAIN: C, D; HIGH MOBILITY GROUP PROTEIN D; CHAIN: A, B;	GENE REGULATION/DNA HMG-D; PROTEIN-DNA COMPLEX, HMG DOMAIN, NON-SEQUENCE SPECIFIC CHROMOSOMAL PROTEIN, HMG-D
301	1ev2	G	613	699	7.2e-07	-0.07	0.11		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGF2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
302	1chc		37	83	1.4e-12	0.02	0.83		VIRUS EQUINE HERPES VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	
302	1fbv	A	39	94	1.4e-12	-0.34	0.52		SIGNAL TRANSDUCTION PROTEIN CBL; CHAIN: A; ZAP-70 PEPTIDE; CHAIN: B; UBIQUITIN-CONJUGATING ENZYME E12-18 KDA UBCH7; CHAIN: C;	LIGASE CBL, UBCH7, ZAP-70, E2, UBIQUITIN, E3, PHOSPHORYLATION, 2 TYROSINE KINASE, UBIQUITINATION, PROTEIN DEGRADATION,
302	1g25	A	35	89	1.5e-05	0.30	0.27		CDK-ACTIVATING KINASE ASSEMBLY FACTOR MAT1; CHAIN: A;	METAL BINDING PROTEIN RING FINGER PROTEIN MAT1; RING FINGER (C3HC4)
302	1rmd		39	117	3.4e-16	0.03	0.65		RAG1; CHAIN: NULL;	DNA-BINDING PROTEIN V(D)J RECOMBINATION ACTIVATING PROTEIN 1; RAG1, V(D)J RECOMBINATION, ANTIBODY, MAD, RING FINGER, 2 ZINC BINUCLEAR CLUSTER, ZINC FINGER, DNA-BINDING PROTEIN
304	1bhl		2	49	6.8e-13	-0.90	0.22		HIV-1 INTEGRASE; CHAIN: NULL;	DNA INTEGRATION DNA INTEGRATION, AIDS, POLYPROTEIN, HYDROLASE, 2 ENDONUCLEASE, POLYNUCLEOTIDYL TRANSFERASE, DNA BINDING 3 (VIRAL)
304	1bl3	C	2	52	1e-13	-0.63	0.06		INTEGRASE; CHAIN: A, B, C;	DNA INTEGRATION DNA INTEGRATION, AIDS, POLYPROTEIN,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
304	1c0m	A	18	97	8.5e-19	0.02	0.81		INTEGRASE; CHAIN: A, B, C, D;	HYDROLASE, 2 ENDONUCLEASE, POLYNUCLEOTIDYL TRANSFERASE, DNA BINDING 3 (VIRAL) TRANSFERASE INTEGRASE, ROUS SARCOMA VIRUS, HIV, X-RAY CRYSTALLOGRAPHY, 2 PROTEIN STRUCTURE, TRANSFERASE
304	1c1a	B	18	97	6.8e-16	-0.28	0.70		RSV INTEGRASE; CHAIN: A, B;	VIRUS/VIRAL PROTEIN INTEGRASE, ROUS SARCOMA VIRUS, HIV, X- RAY CRYSTALLOGRAPHY, 2 VIRUS/VIRAL PROTEIN
304	1ex4	A	2	96	5.1e-33	-0.18	0.62		INTEGRASE; CHAIN: A, B;	VIRUS/VIRAL PROTEIN SH3- LIKE DOMAIN, NONSPECIFIC DNA BINDING BETA SHEET, CIS- 2 PROLINE
304	1hiv	A	51	96	1.2e-18	0.11	0.48		HIV-1 INTEGRASE; CHAIN: NULL	DNA-BINDING PROTEIN DNA-BINDING PROTEIN, AIDS, POLYPROTEIN
304	1qs4	A	2	52	8.5e-13	-0.88	0.15		HIV-1 INTEGRASE; CHAIN: A, B, C;	HYDROLASE DNA INTEGRATION, INTEGRASE) HIV, HYDROLASE, ASPARTYL 2 PROTEASE, ENDONUCLEASE
305	1b7t	A	1	803	0			413.91	MYOSIN HEAVY CHAIN; CHAIN: A; MYOSIN REGULATORY LIGHT CHAIN; CHAIN: Y; MYOSIN ESSENTIAL LIGHT CHAIN; CHAIN: Z;	MYOSIN MYOSIN MOTOR
305	1b7t	A	1	809	0	0.50	1.00		MYOSIN HEAVY CHAIN; CHAIN: A; MYOSIN REGULATORY LIGHT	MYOSIN MYOSIN MOTOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
305	1br1	A	1	775	0	0.56	1.00		CHAIN: Y; MYOSIN ESSENTIAL LIGHT CHAIN; CHAIN: Z; MYOSIN; CHAIN: A, B, C, D, E, F, G, H;	MUSCLE PROTEIN MDE; MUSCLE PROTEIN
305	1br1	A	1	777	0			451.74	MYOSIN; CHAIN: A, B, C, D, E, F, G, H;	MUSCLE PROTEIN MDE; MUSCLE PROTEIN
305	1br2	A	1	746	0	0.57	1.00		MYOSIN; CHAIN: A, B, C, D, E, F;	MUSCLE PROTEIN MUSCLE PROTEIN
305	1br2	A	1	746	0			417.00	MYOSIN; CHAIN: A, B, C, D, E, F;	MUSCLE PROTEIN MUSCLE PROTEIN
305	1dfk	A	1	809	0	0.45	1.00		MYOSIN HEAD; CHAIN: A; MYOSIN HEAD; CHAIN: Y; MYOSIN HEAD; CHAIN: Z;	CONTRACTILE PROTEIN MYOSIN MOTOR, CONFORMATIONAL CHANGES
305	1lvk		2	709	0			420.64	MYOSIN; CHAIN: NULL;	CONTRACTILE PROTEIN MYOSIN, DICTYOSTELIUM, MOTOR, MANT, ATPASE, ACTIN-BINDING, 2 COILED COIL
305	1lvk		3	746	0	0.29	1.00		MYOSIN; CHAIN: NULL;	CONTRACTILE PROTEIN MYOSIN, DICTYOSTELIUM, MOTOR, MANT, ATPASE, ACTIN-BINDING, 2 COILED COIL
305	1mnd		2	644	0			370.75	MYOSIN; CHAIN: NULL;	CONTRACTILE PROTEIN ATPASE, MYOSIN, COILED COIL, ACTIN-BINDING, ATP BINDING, 2 HEPTAD REPEAT PATTERN, METHYLATION, ALKYLATION, 3 PHOSPHORYLATION, CONTRACTILE PROTEIN
305	1mnd		3	644	0	0.71	1.00		MYOSIN; CHAIN: NULL;	CONTRACTILE PROTEIN ATPASE, MYOSIN, COILED COIL, ACTIN-BINDING, ATP BINDING, 2 HEPTAD REPEAT PATTERN,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
305	2mys	A	1	785	0	0.09	1.00		MYOSIN; CHAIN: A, B, C;	METHYLATION, ALKYLATION, 3 PHOSPHORYLATION, CONTRACTILE PROTEIN
305	2mys	A	1	807	0			319.94	MYOSIN; CHAIN: A, B, C;	MUSCLE PROTEIN MUSCLE PROTEIN, MYOSIN SUBFRAGMENT-1, MYOSIN HEAD, 2 MOTOR PROTEIN
306	1alh	A	120	202	5.1e-29			87.27	QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
306	1alh	A	12	116	1.7e-20	-0.19	0.29		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
306	1alh	A	148	229	5.1e-29	0.46	0.86		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
306	1alh	A	176	244	8.5e-24	0.05	0.25		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
306	1alh	A	278	356	3.4e-25	0.29	0.25		QGR ZINC FINGER PEPTIDE; CHAIN: A;	COMPLEX (ZINC FINGER/DNA) COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
306	1ard		176	203	5.1e-05	0.08	0.70		DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C; TRANSCRIPTION REGULATION YEAST TRANSCRIPTION FACTOR ADR1 (RESIDUES 102 - 130) 1ARD 3 (AMINO TERMINAL ZINC FINGER DOMAIN) (NMR, 10 STRUCTURES) 1ARD 4 (ADRI) 1ARD 5 (ADRI) 1ARD 5	(ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
306	1b8t	A	30	127	9.6e-05	-0.34	0.09		CRP1; CHAIN: A;	CONTRACTILE LIM DOMAIN, CRP, NMR, MUSCLE DIFFERENTIATION, CONTRACTILE
306	1mey	C	119	200	1e-50	0.46	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
306	1mey	C	147	229	3.4e-48	0.31	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
306	1mey	C	17	116	1.7e-36	-0.34	0.74		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
306	1mey	C	278	356	5.1e-41	0.42	0.60		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
306	1mey	C	314	382	3.4e-30	0.27	0.12		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
306	1mey	C	66	144	1.7e-48	0.43	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
306	1mey	C	91	172	1e-50	0.29	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
306	1mey	C	91	173	1e-50			111.82	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
306	1paa		176	202	6.8e-05	-0.41	0.60		TRANSCRIPTION REGULATION YEAST TRANSCRIPTION FACTOR ADRI (RESIDUES 130 - 159)	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									IPAA 3 (PAPA - CARBOXY TERMINAL ZINC FINGER DOMAIN) MUTANT WITH IPAA 4 PRO 131 REPLACED BY ALA, PRO 133 REPLACED BY ALA, CYS 140 IPAA 5 REPLACED BY ALA (P131A,P133A,C140A) (NMR, 10 STRUCTURES) IPAA 6	
306	1sp2		176	203	1.7e-07	0.05	0.94		SPIF2; CHAIN: NULL;	ZINC FINGER TRANSCRIPTION FACTOR SPI; ZINC FINGER, TRANSCRIPTION ACTIVATION, SPI
306	1tf3	A	176	237	8.5e-16	0.20	-0.09		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
306	1tf3	A	26	116	1.5e-13	-0.07	0.25		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
306	1tf3	A	278	355	1.2e-15	0.20	-0.02		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
306	1tf3	A	63	147	1.5e-13			60.98	TRANSCRIPTION FACTOR IIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
306	1tf6	A	176	348	1.5e-27	0.15	-0.06		TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIA; 5S GENE; NMR, TFIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
306	1tf6	A	26	181	1.4e-30	0.09	0.75		TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) RNA POLYMERASE II, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
306	1tf6	A	278	380	6.8e-17	0.37	0.12		TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE II, 2 TRANSCRIPTION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
306	1tf6	A	64	224	8.5e-37			108.72	TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	INITIATION, ZINC FINGER PROTEIN
306	1tf6	A	66	210	8.5e-37	0.27	1.00		TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 INITIATION, ZINC FINGER PROTEIN
306	1tf6	A	92	238	5.1e-36	0.03	0.98		TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 INITIATION, ZINC FINGER PROTEIN
306	1ubd	C	127	229	1.5e-32	0.36	0.94		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
306	1ubd	C	155	302	1.4e-26	0.16	-0.14		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YANG 1; TRANSCRIPTION INITIATION, INITIATOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
306	1ubd	C	17	116	1.7e-20	-0.22	0.07		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
306	1ubd	C	280	356	8.5e-23	0.07	0.53		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
306	1ubd	C	62	173	1.7e-34			103.37	YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
306	1ubd	C	71	172	5.1e-34	0.32	1.00		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
306	1ubd	C	99	200	1.7e-34	0.22	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
306	2adr		148	206	6.8e-15			51.79	ADR1; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADR1, ZINC FINGER, NMR
306	2adr		176	231	6.8e-15	-0.35	0.25		ADR1; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADR1, ZINC FINGER, NMR
306	2adr		19	90	8.5e-08	-0.30	0.09		ADR1; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADR1, ZINC FINGER, NMR
306	2adr		278	341	3.4e-14	0.41	0.09		ADR1; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADR1, ZINC FINGER, NMR
306	2adr		315	356	1.7e-12	0.19	-0.01		ADR1; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADR1, ZINC FINGER, NMR
306	2gli	A	183	355	3.4e-21	-0.00	-0.15		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
306	2gli	A	26	143	1.7e-26	0.19	0.48		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
306	2gli	A	63	202	1.2e-63			101.31	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
306	2gli	A	65	200	1.2e-63	0.32	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
306	2gli	A	99	228	1.7e-32	0.12	0.72		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
306	6rxn		346	379	0.0036	0.63	0.96		ELECTRON TRANSFER(IRON-SULFUR PROTEIN) RUBREDOXIN 6RXN 3	
307	1a5f	H	5	216	5.1e-18			50.83	MONOCLONAL ANTI-E-SELECTIN 7A9 ANTIBODY; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB, ANTIBODY, ANTI-E-SELECTIN
307	1afv	H	25	207	1e-18	-0.14	0.98		HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 CAPSID CHAIN: A, B; ANTIBODY FAB25.3 FRAGMENT; CHAIN: H, K, L, M;	COMPLEX (VIRAL CAPSID/IMMUNOGLOBULIN) HIV-1 CA, HIV CA, HIV P24, P24; FAB, FAB LIGHT CHAIN, FAB HEAVY CHAIN, COMPLEX (VIRAL CAPSID/IMMUNOGLOBULIN), HIV, CAPSID PROTEIN, 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
307	1ao7	E	22	214	0.0036			52.66	HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL RECEPTOR ALPHA; CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	P24 COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA-A2 HEAVY CHAIN; CLASS I MHC, T-CELL RECEPTOR, VIRAL PEPTIDE, 2 COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR
307	1aqk	L	13	210	8.5e-16			51.01	FAB B7-15A2; CHAIN: L, H;	IMMUNOGLOBULIN HUMAN FAB, ANTI-TETANUS TOXOID, HIGH AFFINITY, CRYSTAL 2 PACKING MOTIF, PROGRAMMING PROPENSITY TO CRYSTALLIZE, 3 IMMUNOGLOBULIN
307	1ay1	H	5	216	1.5e-15			53.04	TP7 FAB; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, ANTIBODY, FAB, ENZYME INHIBITOR, PCR, 2 HOT START
307	1bql	H	33	207	5.1e-18	0.00	0.18		COMPLEX (ANTIBODY/ANTIGEN) HYHEL-5 FAB COMPLEXED WITH BOBWHITE QUAIL LYSOZYME IBQL 3 IBQL 95	
307	1cf8	H	5	218	3.4e-17			51.50	CATALYTIC ANTIBODY 19A4 (LIGHT CHAIN); CHAIN: L; CATALYTIC ANTIBODY 19A4 (HEAVY CHAIN); CHAIN: H;	CATALYTIC ANTIBODY CATALYTIC ANTIBODY, TERPENOID SYNTHASE, CARBOXYLATION, 2 CYCLIZATION CASCADE
307	1cr9	H	31	211	1.5e-19	0.10	0.30		FAB ANTIBODY LIGHT CHAIN; CHAIN: L; FAB ANTIBODY HEAVY CHAIN; CHAIN: H;	IMMUNE SYSTEM ANTI-PRION FAB 3F4; ANTI-PRION FAB 3F4 ANTI-PRION ANTIBODY, FAB 3F4

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
307	1cs6	A	26	211	3.4e-23	-0.13	0.48		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
307	1cvs	C	37	209	5.1e-17	0.01	0.09		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
307	1dbb	H	25	207	5.1e-19	-0.40	0.05		IMMUNOGLOBULIN FAB' FRAGMENT OF THE DB3 ANTI-STERIOD MONOCLONAL ANTIBODY IDBB 3 (IGG1, SUBGROUP 2A, KAPPA 1) COMPLEX WITH PROGESTERONE IDBB 4	
307	1dgi	R	23	218	1.7e-25	-0.45	0.31		POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PVR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS-RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR
307	1dqg	B	25	207	5.1e-19	-0.02	0.24		ANTI-LYSOZYME ANTIBODY HYHEL-63 (LIGHT CHAIN); CHAIN: A, C; ANTI-LYSOZYME ANTIBODY HYHEL-63 (HEAVY CHAIN); CHAIN: B, D;	IMMUNE SYSTEM ANTI-LYSOZYME ANTIBODY, HYHEL-63, HEN EGG WHITE LYSOZYME
307	1emt	H	25	207	3.4e-19	0.19	0.83		IGG ANTIBODY (LIGHT CHAIN); CHAIN: L; IGG ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM ANTI-FULLERENE ANTIBODY, NANOTUBES
307	1ctz	H	36	212	1.4e-16	0.08	0.69		FAB NC10.14 - LIGHT CHAIN; CHAIN: L, A; FAB NC10.14 - HEAVY CHAIN;	IMMUNE SYSTEM ANTI-SWEETENER FAB, ANTIGEN-ANTIBODY,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									CHAIN: H, B	COMPLEX, CRYSTAL 2 STRUCTURE, RECEPTOR MIMICRY, ANTIGEN RECOGNITION
307	1ev2	G	37	199	5.1e-17	-0.02	0.03		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
307	1evt	C	37	209	1.2e-17	0.08	0.05		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
307	1f11	B	25	207	1e-18	0.20	0.90		F124 IMMUNOGLOBULIN (KAPPA LIGHT CHAIN); CHAIN: A, C; F124 IMMUNOGLOBULIN (IGG1 HEAVY CHAIN); CHAIN: B, D;	IMMUNE SYSTEM IMMUNOGLOBULIN, ANTIBODY, FAB, HEPATITIS B, PRES2
307	1f3d	H	25	207	1.7e-18	0.11	0.57		CATALYTIC ANTIBODY 4B2; CHAIN: L, J; CATALYTIC ANTIBODY 4B2; CHAIN: H, K;	IMMUNE SYSTEM CATALYTIC ANTIBODY, AMIDINIUM, HAPTENIC CHARGE
307	1f3	H	31	207	6.8e-20	-0.12	0.65		BLUE FLUORESCENT ANTIBODY (19G2)-HEAVY CHAIN; CHAIN: H, A; BLUE FLUORESCENT ANTIBODY (19G2)-LIGHT CHAIN; CHAIN: L, B;	IMMUNE SYSTEM IMMUNOGLOBULIN FOLD
307	1f1r	H	31	207	1.7e-18	-0.04	0.42		4-4-20 (IG*G2A=KAPPA=) FAB FRAGMENT; IFLR 5 CHAIN: L, H; IFLR 6	IMMUNOGLOBULIN
307	1gpo	H	25	207	3.4e-20	0.07	0.57		ANTIBODY M41; CHAIN: L, H, M, I;	IMMUNOGLOBULIN PROTEIN ENGINEERING, M

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
307	ligy	B	31	207	3.4e-18	-0.29	0.43		IGG1 INTACT ANTIBODY MAB61.1.3; CHAIN: A, B, C, D	ANTIBODY DESIGN, IMMUNOGLOBULIN 2 STRUCTURE, ANTIGEN-BINDING SITE, CANONICAL CONFORMATION, 3 COMPLEMENTARITY-DETERMINING REGION
307	lind	H	32	207	1.4e-18	0.09	0.51		IMMUNOGLOBULIN CHA255 IMMUNOGLOBULIN FAB' FRAGMENT (IGG1-LAMBDA) COMPLEX 1IND 3 WITH 4-[N'-2-HYDROXYETHYL]-THIOUREIDO]-L-BENZYL-1IND 4 EDTA-IN(3+)(INDIUM(3+)-EOTUBE) 1IND 5	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN, V REGION, C REGION, HINGE REGION
307	lmlb	B	33	207	3.4e-17	0.31	0.45		IMMUNOGLOBULIN FAB D44.1 (IGG1.KAPPA) (BALB/C MOUSE, MONOCLONAL ANTIBODY) 1MLB 5	
307	lnsn	H	25	207	5.1e-19	0.03	0.84		IGG FAB (IGG1, KAPPA); INSN 4 CHAIN: L, H; INSN 5 STAPHYLOCOCCAL NUCLEASE; INSN 9 CHAIN: S; INSN 10	COMPLEX (IMMUNOGLOBULIN/HYDROLASE) N10 FAB IMMUNOGLOBULIN; INSN STAPHYLOCOCCAL NUCLEASE RIBONUCLEASE, INSN 11 IMMUNOGLOBULIN, STAPHYLOCOCCAL NUCLEASE INSN 25
307	lsbs	H	31	207	1.7e-19	-0.07	0.54		MONOCLONAL ANTIBODY 3A2; CHAIN: H,	MONOCLONAL ANTIBODY

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									L;	FAB-FRAGMENT, REPRODUCTION
307	1sm3	H	31	207	1e-20	-0.03	0.94		SM3 ANTIBODY; CHAIN: L, H; PEPTIDE EPITOPE; CHAIN: P;	COMPLEX (ANTIBODY/PEPTIDE EPITOPE) ANTIBODY, PEPTIDE ANTIGEN, ANTITUMOR ANTIBODY, 2 COMPLEX (ANTIBODY/PEPTIDE EPITOPE)
307	1sm3	L	22	206	5.1e-15			53.74	SM3 ANTIBODY; CHAIN: L, H; PEPTIDE EPITOPE; CHAIN: P;	COMPLEX (ANTIBODY/PEPTIDE EPITOPE) ANTIBODY, PEPTIDE ANTIGEN, ANTITUMOR ANTIBODY, 2 COMPLEX (ANTIBODY/PEPTIDE EPITOPE)
307	1tet	H	25	207	3.4e-19	-0.32	0.30		IMMUNOGLOBULIN IGG1 MONOCLONAL FAB FRAGMENT (TE33) COMPLEX WITH CHOLERA 1TET 3 TOXIN PEPTIDE 3 (CTP3) 1TET 4	
307	25c8	H	31	207	1.7e-18	-0.18	0.82		IGG 5C8; CHAIN: L, H;	CATALYTIC ANTIBODY CATALYTIC ANTIBODY, FAB, RING CLOSURE REACTION
307	2cgr	H	33	207	1.7e-16	0.28	0.18		IMMUNOGLOBULIN IGG2B (KAPPA) FAB FRAGMENT COMPLEXED WITH ANTIGEN 2CGR 3 N-(P-CYANOPHENYL)-N-(DIPHENYLEMETHYL) GUANIDINEACETIC ACID 2CGR 4	
307	3hfm	H	25	207	1e-20	-0.16	0.39		COMPLEX(ANTIBODY-ANTIGEN) IG*G1 FAB FRAGMENT (HY/HEL5-10)	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									AND LYSOZYME (E.C.3.2.1.17) 3HFM 4 COMPLEX 3HFM 5	
309	1cly	A	1	163	1.2e-63			57.43	RAS-RELATED PROTEIN RAP-1A; CHAIN: A; PROTO-ONCOGENE SERINE/THREONINE PROTEIN KINASE CHAIN: B;	SIGNALING PROTEIN GTP-BINDING PROTEINS, PROTEIN-PROTEIN COMPLEX, EFFECTORS
309	1cly	A	3	163	1.2e-63	0.28	0.45		RAS-RELATED PROTEIN RAP-1A; CHAIN: A; PROTO-ONCOGENE SERINE/THREONINE PROTEIN KINASE CHAIN: B;	SIGNALING PROTEIN GTP-BINDING PROTEINS, PROTEIN-PROTEIN COMPLEX, EFFECTORS
309	1cc0	A	4	171	1.2e-51	0.04	0.00		TRANSFORMING PROTEIN RHOA; CHAIN: A, C; RHO GDP DISSOCIATION INHIBITOR ALPHA; CHAIN: E, F;	SIGNALING PROTEIN GTP-BINDING PROTEIN RHOA, GTPASE RHOA; RHO GDI 1; RHO GTPASE, G-PROTEIN, SIGNALING PROTEIN
309	1ctq	A	1	164	1e-65			51.73	TRANSFORMING PROTEIN P21/H-RAS-1; CHAIN: A;	SIGNALING PROTEIN G PROTEIN, GTP HYDROLYSIS, KINETIC CRYSTALLOGRAPHY, 2 SIGNALING PROTEIN
309	1ctq	A	3	163	1e-65	0.34	0.36		TRANSFORMING PROTEIN P21/H-RAS-1; CHAIN: A;	SIGNALING PROTEIN G PROTEIN, GTP HYDROLYSIS, KINETIC CRYSTALLOGRAPHY, 2 SIGNALING PROTEIN
309	1cxz	A	4	163	5.1e-53	0.34	0.16		HIS-TAGGED TRANSFORMING PROTEIN RHOA(0-181); CHAIN: A; PKN; CHAIN: B;	SIGNALING PROTEIN PROTEIN-PROTEIN COMPLEX, ANTIPARALLEL COILED-COIL
309	1d5c	A	3	161	3.4e-54	0.30	0.25		RAB6 GTPASE; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS G-PROTEIN, GTPASE, RAB6, VESICULAR TRAFFICKING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
309	1ek0	A	2	164	3.4e-51	0.41	0.29		GTP-BINDING PROTEIN YPT51; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS G PROTEIN, VESICULAR TRAFFIC, GTP HYDROLYSIS, YPT/RAB 2 PROTEIN, ENDOCYTOSIS, HYDROLASE
309	1kao		1	164	6.8e-60			68.65	RAP2A; CHAIN: NULL;	GTP-BINDING PROTEIN GTP-BINDING PROTEIN, SMALL G PROTEIN, RAP2, GDP, RAS
309	1kao		3	160	6.8e-60	0.14	0.72		RAP2A; CHAIN: NULL;	GTP-BINDING PROTEIN GTP-BINDING PROTEIN, SMALL G PROTEIN, RAP2, GDP, RAS
309	1mh1		3	166	3.4e-52	0.10	-0.01		RAC1; CHAIN: NULL;	GTP-BINDING GTP-BINDING, GTPASE, SMALL G-PROTEIN, RHO FAMILY, RAS SUPER 2 FAMILY
309	1zbd	A	3	167	6.8e-61	0.05	0.34		RAB-3A; CHAIN: A; RABPHILIN-3A; CHAIN: B;	COMPLEX (GTP-BINDING/EFFECTOR) RAS-RELATED PROTEIN RAB3A; COMPLEX (GTP-BINDING/EFFECTOR), G PROTEIN, EFFECTOR, RABCDR, 2 SYNAPTIC EXOCYTOSIS, RAB PROTEIN, RAB3A, RABPHILIN
309	3rab	A	3	164	1.7e-61	-0.08	0.36		RAB3A; CHAIN: A;	HYDROLASE G PROTEIN, VESICULAR TRAFFICKING, GTP HYDROLYSIS, RAB 2 PROTEIN, NEUROTRANSMITTER RELEASE, HYDROLASE
310	1ao7	D	23	118	2.4e-32			73.76	HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA-A2 HEAVY CHAIN; CLASS II MHC, T-CELL RECEPTOR, T

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
310	1b0w	A	23	118	5.1e-43			57.66	RECEPTOR ALPHA; CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	VIRAL PEPTIDE, 2 COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR
310	1b6d	A	24	113	1.7e-44	0.21	0.95		BENCE-JONES KAPPA I PROTEIN BRE; CHAIN: A, B, C;	IMMUNE SYSTEM BENCE-JONES; IMMUNOGLOBULIN AMYLOID, IMMUNE SYSTEM
310	1b88	A	23	118	8.5e-40			87.03	IMMUNOGLOBULIN; CHAIN: A, B;	IMMUNOGLOBULIN LIGHT-CHAIN DIMER HEADER
310	1bd2	D	23	118	1e-34			55.59	T CELL RECEPTOR V-ALPHA DOMAIN; CHAIN: A, B;	T CELL RECEPTOR TCR; T CELL RECEPTOR, MHC CLASS I, HUMAN IMMUNODEFICIENCY VIRUS, 2 MOLECULAR RECOGNITION
310	1bj1	L	24	115	3.4e-46	0.24	0.99		HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL RECEPTOR ALPHA; CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA A2 HEAVY CHAIN; COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR)
310	1bw	A	20	118	1.2e-43			54.30	FAB FRAGMENT; CHAIN: L, H, J, K; VASCULAR ENDOTHELIAL GROWTH FACTOR; CHAIN: V, W;	COMPLEX (ANTIBODY/ANTIGEN) FAB-12; VEGF; COMPLEX (ANTIBODY/ANTIGEN), ANGIOGENIC FACTOR
310	1bw	A	24	113	1.2e-43	0.51	0.99		IG KAPPA CHAIN V-I REGION REI; CHAIN: A, B;	IMMUNE SYSTEM REIV, STABILIZED IMMUNOGLOBULIN FRAGMENT, BENCE-JONES 2 PROTEIN, IMMUNE SYSTEM
310	1bw	A	24	113	1.2e-43	0.51	0.99		IG KAPPA CHAIN V-I REGION REI; CHAIN: A, B;	IMMUNE SYSTEM REIV, STABILIZED

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
310	1dec	A	24	113	1e-47	0.12	0.98		IGM RF 2A2; CHAIN: A, C, E; IGM RF 2A2; CHAIN: B, D, F; IMMUNOGLOBULIN G BINDING PROTEIN A; CHAIN: G, H;	IMMUNOGLOBULIN FRAGMENT, BENCE-JONES 2 PROTEIN, IMMUNE SYSTEM
310	1dfb	L	24	113	3.4e-46	0.43	1.00		IMMUNOGLOBULIN 3D6 FAB 1DFB 3	IMMUNE SYSTEM FAB-IBP COMPLEX CRYSTAL STRUCTURE 2.7 Å
310	1dql	L	24	113	5.1e-44	0.66	0.99		IGM MEZ IMMUNOGLOBULIN; CHAIN: L; IGM MEZ IMMUNOGLOBULIN; CHAIN: H;	RESOLUTION BINDING 2 OUTSIDE THE ANTIGEN COMBINING SITE
310	1fgv	L	23	117	1.7e-46			55.38	IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 1FGV 3 ANTIBODY 'H52' (HUH52-AA FV) 1FGV 4	SUPERANTIGEN FAB VH3 3 SPECIFICITY
310	1fgv	L	24	113	1.7e-46	0.37	0.99		IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 1FGV 3 ANTIBODY 'H52' (HUH52-AA FV) 1FGV 4	
310	1igm	L	23	118	1.7e-44			52.10	IMMUNOGLOBULIN IMMUNOGLOBULIN M (IG-M) FV FRAGMENT 1IGM 3	
310	1igm	L	24	113	1.7e-44	0.42	1.00		IMMUNOGLOBULIN IMMUNOGLOBULIN M (IG-M) FV FRAGMENT 1IGM 3	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
310	1kb5	A	23	118	5.1e-42			84.21	KB5-C20 T-CELL ANTIGEN RECEPTOR; CHAIN: A, B; ANTIBODY DESIRE-1; CHAIN: L, H;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) TCR VAPLHA VBETA DOMAIN: T-CELL RECEPTOR, STRAND SWITCH, FAB, ANTICLONOTYPIC, 2 (IMMUNOGLOBULIN/RECEPTOR)
310	1vge	L	25	113	1.5e-43	0.55	0.96		TR1.9 FAB; CHAIN: L, H;	IMMUNOGLOBULIN TR1.9, ANTI-THYROID PEROXIDASE, AUTOANTIBODY, 2 IMMUNOGLOBULIN
310	1wtl	A	23	118	1.7e-42			50.36	IMMUNOGLOBULIN WAT, A VARIABLE DOMAIN FROM IMMUNOGLOBULIN LIGHT-CHAIN 1WTL 3 (BENCE-JONES PROTEIN) 1WTL 4	
310	2fgw	L	24	113	3.4e-46	0.43	0.93		IMMUNOGLOBULIN FAB FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 2FGW 3 ANTIBODY 'H52' (HUH52-OZ.FAB) 2FGW 4	
311	1aw9		1	78	5.1e-14	0.01	0.04		GLUTATHIONE S-TRANSFERASE III; CHAIN: NULL;	TRANSFERASE TRANSFERASE, HERBICIDE DETOXIFICATION
311	1axd	A	1	68	1.4e-14	0.17	-0.09		GLUTATHIONE S-TRANSFERASE I; CHAIN: A, B; LACTOYLGLUTATHIONE; CHAIN: C, D	COMPLEX (TRANSFERASE/LIGAND) COMPLEX (TRANSFERASE/LIGAND), TRANSFERASE, HERBICIDE 2 DETOXIFICATION HEADER
311	1b64		135	225	1.4e-46			137.96	ELONGATION FACTOR 1-	GUANINE NUCLEOTIDE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									BETA; CHAIN: NULL;	EXCHANGE FACTOR GUANINE NUCLEOTIDE EXCHANGE FACTOR, G- PROTEIN, TRANSLATION 2 ELONGATION
311	1b64		142	225	1.4e-46	0.43	1.00		ELONGATION FACTOR 1- BETA; CHAIN: NULL;	GUANINE NUCLEOTIDE EXCHANGE FACTOR GUANINE NUCLEOTIDE EXCHANGE FACTOR, G- PROTEIN, TRANSLATION 2 ELONGATION
311	1b8x	A	5	91	5.1e-13	0.24	0.27		AML-1B; CHAIN: A;	SIGNAL PROTEIN NUCLEAR MATRIX TARGETING SIGNAL PROTEIN
311	1bg5		5	91	5.1e-13	0.28	0.33		FUSION PROTEIN OF ALPHA-NA,K-ATPASE WITH CHAIN: NULL;	ANKYRIN BINDING MAB; ANKYRIN BINDING, ATPASE, GLUTATHIONE-S- TRANSFERASE, CARRIER 2 CRYSTALLIZATION, ION TRANSPORT
311	1dug	A	5	91	5.1e-13	0.19	0.27		CHIMERA OF GLUTATHIONE S- TRANSFERASE- SYNTHETIC CHAIN: A, B;	TRANSFERASE, BLOOD CLOTTING GAMMA CHAIN INTEGRIN FRAGMENT, CARRIER PROTEIN DRIVEN 2 CRYSTALLIZATION
311	1f60	B	137	225	4.8e-48	0.20	1.00		ELONGATION FACTOR BEEF1A; CHAIN: A; ELONGATION FACTOR BEEF1B; CHAIN: B;	TRANSLATION PROTEIN- PROTEIN COMPLEX
311	1fhe		2	92	3.4e-16	0.07	0.31		GLUTATHIONE TRANSFERASE; CHAIN: NULL;	TRANSFERASE GST; TRANSFERASE, DETOXIFICATION, GLUTATHIONE TRANSFERASE
311	1gne		5	91	5.1e-13	0.26	0.12		GLUTATHIONE TRANSFERASE GLUTATHIONE S- TRANSFERASE (B.C.2.5.1.18) FUSED WITH	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
311	1gta		5	91	5.1e-13	-0.04	0.24		A 1GNE 3 CONSERVED NEUTRALIZING EPITOPE ON GP41 OF HUMAN 1GNE 4 IMMUNODEFICIENCY VIRUS TYPE 1, COMPLEXED WITH GLUTATHIONE 1GNE 5	
311	1gtu	A	3	86	5.1e-15	0.15	-0.01		GLUTATHIONE TRANSFERASE S-GLUTATHIONE S-TRANSFERASE (E.C.2.5.1.18) (26 KDA) 1GTA 3	TRANSFERASE TRANSFERASE, GLUTATHIONE, CONJUGATION, DETOXIFICATION, 2 CYTOSOLIC, DIMER
312	1buo	A	40	164	1.4e-25	0.19	0.98		PROMYELOCYTIC LEUKEMIA ZINC FINGER PROTEIN PLZF; CHAIN: A;	GENE REGULATION POZ DOMAIN; PROTEIN-PROTEIN INTERACTION DOMAIN, TRANSCRIPTIONAL 2 REPRESSOR, ZINC-FINGER PROTEIN, X-RAY CRYSTALLOGRAPHY, 3 PROTEIN STRUCTURE, PROMYELOCYTIC LEUKEMIA, GENE REGULATION
312	1gof		365	615	1.7e-12	0.25	0.01		OXIDOREDUCTASE(OXYGEN(A)) GALACTOSE OXIDASE (E.C.1.1.3.9) (PH 4.5) 1GOF 3	
313	1alh	A	118	220	2.4e-15	-0.01	0.36		QGSR ZINC FINGER PEPTIDE; CHAIN: A;	COMPLEX (ZINC FINGER/DNA) COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
313	1alh	A	144	219	3.4e-24	0.03	-0.13		DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	(ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
313	1alh	A	167	247	5.1e-28	-0.26	0.17		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
313	1alh	A	195	303	8.5e-24	-0.03	0.53		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
313	1alh	A	256	332	8.4e-30	-0.02	0.74		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
313	1alh	A	279	359	6e-35	0.17	1.00		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
313	1alh	A	307	387	6.8e-31	0.28	1.00		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
313	1alh	A	307	388	1.2e-39	0.35	0.99		C; QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
313	1alh	A	363	472	6e-39	-0.15	0.87		C; QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
313	1alh	A	391	473	6e-39			76.59	C; QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
313	1bbo		254	307	1.1e-15	0.01	0.06		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11 IBBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) IBBO 4	
313	1mey	C	143	219	1.7e-41	0.18	0.15		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
313	1mey	C	166	247	3.4e-47	-0.23	0.52		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
313	1mey	C	222	331	6.8e-45	0.27	0.46		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	(ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
313	1mey	C	278	359	1.7e-50	0.31	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
313	1mey	C	306	387	3.4e-51	0.20	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
313	1mey	C	334	415	3.4e-51	0.22	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
313	1mey	C	362	443	8.5e-51	0.24	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
313	1mey	C	390	471	5.1e-51	0.30	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
313	1mev	C	390	472	5.1e-51			90.72	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
313	1mev	C	418	484	3.4e-42	0.09	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
313	1mev	G	304	331	3.4e-13	0.03	0.96		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
313	1rf3	A	195	303	1e-15	-0.31	0.45		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
313	1rf3	A	196	305	7.2e-18	-0.17	0.04		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
313	1tf3	A	258	333	2.4e-19	-0.04	0.25		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
313	1tf6	A	167	312	5.1e-31	-0.29	0.05		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
313	1tf6	A	268	396	8.5e-35	-0.11	0.82		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
313	1tf6	A	279	431	8.5e-38	0.11	0.98		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
313	1tf6	A	334	485	1.7e-39			94.43	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	PROTEIN COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
313	1tf6	A	335	485	1.7e-39	0.37	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
313	1tf6	A	86	228	8.5e-32	-0.37	0.01		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
313	1ubd	C	154	247	5.1e-29	-0.23	0.21		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
313	1ubd	C	168	303	3.4e-28	-0.17	0.00		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
313	1ubd	C	262	359	1.5e-33	-0.12	0.98		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
313	1ubd	C	286	387	1e-34	0.06	0.95		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
313	1ubd	C	314	415	6.8e-35	0.02	1.00		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
313	1ubd	C	362	472	3.4e-35			79.23	YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
313	1ubd	C	370	471	1e-34	-0.19	1.00		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	REGULATION/DNA COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
313	1ubd	C	398	484	5.1e-29	-0.16	1.00		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
313	2gli	A	116	305	4.8e-29	-0.16	0.43		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
313	2gli	A	166	330	5.1e-30	-0.31	0.05		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
313	2gli	A	196	417	2.4e-53	-0.32	0.52		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
313	2gli	A	251	389	6e-54	0.12	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
313	2gli	A	262	386	8.5e-34	0.19	0.98		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
313	2gli	A	307	473	9.6e-64	0.22	0.98		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
313	2gli	A	308	445	9.6e-64			88.39	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
313	2gli	A	342	470	3.4e-34	-0.01	0.98		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
313	2gli	A	363	480	4.8e-56	-0.10	0.99		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
313	2gli	A	370	480	6.8e-30	-0.10	0.87		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
315	Irgs		363	501	2.6e-26	0.36	0.81		CAMP DEPENDENT PROTEIN KINASE; CHAIN: NULL;	KINASE RI(αPHA); REGULATORY SUBUNIT, KINASE
315	2cgp	A	383	504	2.6e-22	0.37	0.35		CATABOLITE GENE ACTIVATOR PROTEIN; CHAIN: A; DNA (5'-D(*GP*TP*CP*AP*CP*AP*TP*TP*AP*AP*T)-3');)	TRANSCRIPTION/DNA COMPLEX (TRANSCRIPTION REGULATION/DNA), DNA-BINDING, CAMP-2 BINDING, ACTIVATOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									CHAIN: B; DNA (5'- CHAIN: C;	
316	1b12	A	38	105	2.6e-15	-0.24	0.78		SIGNAL PEPTIDASE I; CHAIN: A, B, C, D;	HYDROLASE SPASE I, LEADER PEPTIDASE I; SERINE PROTEINASE, SERINE-DEPENDANT HYDROLASE, SIGNAL 2 PEPTIDE PROCESSING, PROTEIN TRANSLOCATION, MEMBRANE BOUND 3 PROTEINASE, MEMBRANE PROTEIN
316	1b12	C	38	127	1.6e-16	-0.25	0.47		SIGNAL PEPTIDASE I; CHAIN: A, B, C, D;	HYDROLASE SPASE I, LEADER PEPTIDASE I; SERINE PROTEINASE, SERINE-DEPENDANT HYDROLASE, SIGNAL 2 PEPTIDE PROCESSING, PROTEIN TRANSLOCATION, MEMBRANE BOUND 3 PROTEINASE, MEMBRANE PROTEIN
316	1b12	D	38	135	1.6e-20	-0.60	0.43		SIGNAL PEPTIDASE I; CHAIN: A, B, C, D;	HYDROLASE SPASE I, LEADER PEPTIDASE I; SERINE PROTEINASE, SERINE-DEPENDANT HYDROLASE, SIGNAL 2 PEPTIDE PROCESSING, PROTEIN TRANSLOCATION, MEMBRANE BOUND 3 PROTEINASE, MEMBRANE PROTEIN
317	1a1h	A	113	206	1.5e-27	0.15	0.41		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
317	1mey	C	125	206	1.5e-44	0.08	0.87		C; DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
317	1mey	C	153	234	3.4e-49	0.48	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
317	1mey	C	181	262	6.8e-50	0.43	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
317	1mey	C	209	290	1.2e-50	0.16	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
317	1mey	C	237	318	1.2e-50	0.36	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
317	1mey	C	265	346	1.2e-50	0.36	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
317	1mey	C	293	374	1e-50	0.50	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
317	1mey	C	293	375	1e-50			93.81	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
317	1mey	C	321	400	1.4e-47	0.45	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
317	1mey	G	151	178	1.7e-12	0.01	0.95		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
317	1tf6	A	154	299	1.4e-37	-0.14	0.99		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER

Table 5

SEQ ID NO;	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
317	1tf6	A	209	370	3.4e-38			99.41	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	PROTEIN COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE II, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
317	1tf6	A	238	382	3.4e-38	0.03	0.86		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE II, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
317	1ubd	C	127	234	5.1e-34	-0.01	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA- PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
317	1ubd	C	158	262	9.9e-50	0.32	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA- PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
317	1ubd	C	161	262	3.4e-35	0.23	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
317	1ubd	C	179	290	2e-50	0.03	1.00		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
317	1ubd	C	217	318	1.7e-34	0.18	1.00		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
317	1ubd	C	245	346	1.7e-35	0.10	1.00		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
317	1ubd	C	263	374	9.9e-51	0.11	1.00		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3

Table 5

SEQ ID No:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
317	1ubd	C	267	375	9.9e-51			91.94	YY1; CHAIN: C; ADENO- ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA) YING- YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA- PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
317	1ubd	C	298	400	3.3e-47	-0.12	0.99		YY1; CHAIN: C; ADENO- ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING- YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA- PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
317	1ubd	C	301	400	1e-31	-0.25	0.86		YY1; CHAIN: C; ADENO- ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING- YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA- PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
317	2drp	A	150	206	9.9e-20	-0.12	0.12		COMPLEX(TRANSCRIPTIO N REGULATION/DNA) TRAMTRACK PROTEIN (TWO ZINC-FINGER PEPTIDE) COMPLEXED WITH 2DRP 3 DNA 2DRP 4	
317	2gli	A	133	261	6.8e-35	-0.00	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE- FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
317	2gli	A	189	317	1.7e-34	0.03	0.47		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	BINDING PROTEIN/DNA) COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
317	2gli	A	209	348	1.7e-34			81.75	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
317	2gli	A	245	373	6.8e-34	0.23	0.87		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
317	2gli	A	273	399	1.7e-30	-0.16	0.86		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
317	2gli	A	65	205	1.7e-30	0.21	-0.12		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
320	1apo		23	59	6.6e-12	0.72	1.00		COAGULATION FACTOR EGF-LIKE MODULE OF BLOOD COAGULATION FACTOR X (N-TERMINAL, 1 APO 3 APO FORM) (NMR, 13 STRUCTURES) 1 APO 4	
320	1bf9		23	61	1.3e-11	0.68	1.00		FACTOR VII; CHAIN: NULL;	BLOOD COAGULATION BLOOD COAGULATION, EGF, HYDROLASE, SERINE PROTEASE
320	1cej	A	24	78	2.3e-09	0.15	-0.13		MEROZOITE SURFACE PROTEIN 1; CHAIN: A;	SURFACE PROTEIN MEROZOITE SURFACE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
320	1dan	L	23	78	3.3e-10	0.30	0.96		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFRCMK) WITH CHAIN: C;	ANTIGEN 1, MAJOR BLOOD-STAGE EGF-LIKE DOMAIN, EXTRACELLULAR, MODULAR PROTEIN, SURFACE 2 ANTIGEN, MALARIA VACCINE COMPONENT, SURFACE PROTEIN
320	legf		26	59	2.6e-09	0.18	0.40		GROWTH FACTOR EPIDERMAL GROWTH FACTOR (EGF) (NMR, 16 STRUCTURES) IEGF 3	BLOOD CLOTTING FACTOR VII, BLOOD COAGULATION, EGF-LIKE DOMAIN, BLOOD 2 CLOTTING
320	17fe	A	23	69	1.3e-11	0.67	0.77		BLOOD COAGULATION FACTOR VII; CHAIN: A;	BLOOD CLOTTING FACTOR VII, BLOOD COAGULATION, EGF-LIKE DOMAIN, BLOOD 2 CLOTTING
320	1fak	L	23	78	3.3e-10	0.42	0.64		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIG AND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIG AND), BLOOD CLOTTING
320	1tpg		24	59	3e-11	0.45	0.99		T-PLASMINOGEN ACTIVATOR F1-G; 1TPG 7 CHAIN: NULL; 1TPG 8	PLASMINOGEN ACTIVATION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
320	1whe		23	59	3.3e-10	0.92	0.98		COAGULATION FACTOR X; CHAIN: NULL;	GLYCOPROTEIN, GLYCOPROTEIN, HYDROLASE, SERINE PROTEASE, PLASMA, BLOOD 2 COAGULATION FACTOR
320	1yuf		26	66	2.3e-09	0.50	-0.07		TRANSFORMING GROWTH FACTOR ALPHA; CHAIN: NULL;	GROWTH FACTOR TGF-ALPHA, H-TGF-ALPHA; EGF-LIKE DOMAIN STRUCTURE, GROWTH FACTOR
321	1b08	A	242	372	2.3e-22	0.69	0.93		LUNG SURFACTANT PROTEIN D; CHAIN: A, B, C;	SUGAR BINDING PROTEIN C-TYPE LECTIN, CRD, SP-D, COLECTIN, ALPHA-HELICAL COILED-2 COIL, LUNG SURFACTANT, SUGAR BINDING PROTEIN
321	1b6e		249	376	1.6e-23	0.32	1.00		CD94; CHAIN: NULL;	NK CELL NK CELL, RECEPTOR, C-TYPE LECTIN, C-TYPE LECTIN-LIKE, NKD
321	1b6e		249	376	1.6e-23			67.41	CD94; CHAIN: NULL;	NK CELL NK CELL, RECEPTOR, C-TYPE LECTIN, C-TYPE LECTIN-LIKE, NKD
321	1bj3	A	249	374	9.9e-20			61.40	COAGULATION FACTOR IX-BINDING PROTEIN A; CHAIN: A; COAGULATION FACTOR IX-BINDING PROTEIN B; CHAIN: B;	COLLAGEN BINDING PROTEIN IX-BP; IX-BP; COAGULATION FACTOR IX-BINDING, HETERODIMER, VENOM, HABU 2 SNAKE, C-TYPE LECTIN SUPERFAMILY, COLLAGEN BINDING PROTEIN
321	1bj3	A	250	372	9.9e-20	0.42	0.99		COAGULATION FACTOR IX-BINDING PROTEIN A; CHAIN: A; COAGULATION FACTOR IX-BINDING PROTEIN B; CHAIN: B;	COLLAGEN BINDING PROTEIN IX-BP; IX-BP; COAGULATION FACTOR IX-BINDING, HETERODIMER, VENOM, HABU 2 SNAKE, C-TYPE LECTIN SUPERFAMILY, COLLAGEN BINDING PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
321	1c3a	A	247	374	6.6e-21	0.61	0.81		FLAVOCETIN-A: ALPHA SUBUNIT; CHAIN: A; FLAVOCETIN-A: BETA SUBUNIT; CHAIN: B	MEMBRANE PROTEIN C-TYPE LECTIN-LIKE DOMAINS
321	1c3a	B	250	375	6.6e-22	0.36	0.98		FLAVOCETIN-A: ALPHA SUBUNIT; CHAIN: A; FLAVOCETIN-A: BETA SUBUNIT; CHAIN: B	MEMBRANE PROTEIN C-TYPE LECTIN-LIKE DOMAINS
321	1e87	A	249	376	1.6e-21	0.91	1.00		EARLY ACTIVATION ANTIGEN CD69; CHAIN: A;	HEMATOPOIETIC CELL RECEPTOR ACTIVATION INDUCER MOLECULE (AIM), EA 1, HEMATOPOIETIC CELL RECEPTOR, LEUCOCYTE, C-TYPE LECTIN-LIKE, 2 NKD, KLR
321	1egg	B	250	375	6.6e-22	0.71	0.75		MACROPHAGE MANNOSE RECEPTOR; CHAIN: A, B;	SUGAR BINDING PROTEIN C-TYPE LECTIN, MANNOSE RECEPTOR
321	1ixx	A	249	374	2e-20			59.78	COAGULATION FACTORS IX/X-BINDING PROTEIN; CHAIN: A, B, C, D, E, F;	COAGULATION FACTOR BINDING IX/X-BP COAGULATION FACTOR BINDING, C-TYPE LECTIN, GLA-DOMAIN 2 BINDING, C-TYPE CRD MOTIF, LOOP EXCHANGED DIMER
321	1ixx	A	250	372	2e-20	0.38	1.00		COAGULATION FACTORS IX/X-BINDING PROTEIN; CHAIN: A, B, C, D, E, F;	COAGULATION FACTOR BINDING IX/X-BP COAGULATION FACTOR BINDING, C-TYPE LECTIN, GLA-DOMAIN 2 BINDING, C-TYPE CRD MOTIF, LOOP EXCHANGED DIMER
321	1ixx	B	249	376	6.6e-23			70.03	COAGULATION FACTORS IX/X-BINDING PROTEIN; CHAIN: A, B, C, D, E, F;	COAGULATION FACTOR BINDING IX/X-BP COAGULATION FACTOR BINDING, C-TYPE LECTIN, GLA-DOMAIN 2 BINDING, C-TYPE CRD MOTIF, LOOP EXCHANGED DIMER

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
321	1lxx	B	250	375	6.6e-23	0.62	0.63		COAGULATION FACTORS IX/X-BINDING PROTEIN; CHAIN: A, B, C, D, E, F;	COAGULATION FACTOR BINDING IX/X-BP COAGULATION FACTOR BINDING, C-TYPE LECTIN, GLA-DOMAIN 2 BINDING, C-TYPE CRD MOTIF, LOOP EXCHANGED DIMER
321	1rdl	1	262	375	1.3e-18			53.97	MANNOSE-BINDING PROTEIN-C; IRDL 6 CHAIN: 1, 2; IRDL 7	LECTIN SUB-MBP-C; IRDL 9 C-TYPE LECTIN, CALCIUM-BINDING PROTEIN IRDL 20
321	1tm3		245	376	6.6e-22			62.26	TETRALECTIN; CHAIN: NULL;	LECTIN TETRALECTIN, PLASMINOGEN BINDING, KRINGLE 4, C-TYPE LECTIN, 2 CARBOHYDRATE RECOGNITION DOMAIN
321	1tm3		249	376	6.6e-22	0.86	1.00		TETRALECTIN; CHAIN: NULL;	LECTIN TETRALECTIN, PLASMINOGEN BINDING, KRINGLE 4, C-TYPE LECTIN, 2 CARBOHYDRATE RECOGNITION DOMAIN
321	2msb	A	262	375	6.6e-19			55.49	LECTIN MANNOSE-BINDING PROTEIN A (LECTIN DOMAIN) COMPLEX WITH 2MSB 3 CALCIUM AND GLYCOPEPTIDE 2MSB 4	
325	1bth	A	24	240	9.9e-20	-0.03	0.04		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
325	1bth	A	36	215	1.3e-23	0.13	0.78		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
325	1cvs	C	23	211	1.3e-18	-0.03	0.70		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR PGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
325	1cvs	D	13	120	3.3e-12	-0.17	0.18		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	FACTOR/GROWTH FACTOR RECEPTOR GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
325	1cvs	D	23	215	1.7e-22	-0.15	0.39		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
325	1cvt	C	23	215	3.3e-20	-0.19	0.18		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	FACTOR/GROWTH FACTOR RECEPTOR FGF1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
325	1fcg	A	24	208	3.3e-20	-0.11	0.24		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32, FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
325	1fhg	A	23	121	9.9e-12	0.17	0.64		TELOKIN; CHAIN: A	CONTRACTILE PROTEIN IMMUNOGLOBULIN FOLD, BETA BARREL
325	1nct		19	121	2.3e-13	0.27	0.24		TTTN; CHAIN: NULL;	MUSCLE PROTEIN CONNECTIN, NEXTM5; CELL ADHESION, GLYCOPROTEIN, TRANSMEMBRANE, REPEAT, BRAIN, 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
325	1tmm		36	121	3.3e-11	0.15	0.17		MUSCLE PROTEIN TITIN MODULE M5 (CONNECTIN) 1TNM 3 (NMR, MINIMIZED AVERAGE STRUCTURE) 1TNM 4 1TNM 58	IMMUNOGLOBULIN FOLD, ALTERNATIVE SPLICING, SIGNAL, 3 MUSCLE PROTEIN
325	1wio	A	33	223	9.9e-19	0.26	0.59		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC LIPOPROTEIN, POLYMORPHISM
325	1wit		24	121	6.6e-14	0.35	0.13		TWITCHIN 18TH IGSF MODULE; CHAIN: NULL;	MUSCLE PROTEIN IMMUNOGLOBULIN SUPERFAMILY, I SET, MUSCLE PROTEIN
325	2dli	A	14	120	1.3e-10	-0.13	0.24		MHC CLASS I NK CELL RECEPTOR PRECURSOR; CHAIN: A;	IMMUNE SYSTEM P58 NATURAL KILLER CELL RECEPTOR; KIR, NATURAL KILLER RECEPTOR, INHIBITORY RECEPTOR, 2 IMMUNOGLOBULIN
325	2dli	A	24	209	6.6e-22	0.07	0.36		MHC CLASS I NK CELL RECEPTOR PRECURSOR; CHAIN: A;	IMMUNE SYSTEM P58 NATURAL KILLER CELL RECEPTOR; KIR, NATURAL KILLER RECEPTOR, INHIBITORY RECEPTOR, 2 IMMUNOGLOBULIN
325	2feb	A	24	209	3.3e-24	-0.03	0.21		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
325	3ncm	A	25	120	6.6e-11	0.46	0.12		NEURAL CELL ADHESION MOLECULE, LARGE ISOFORM; CHAIN: A;	CELL ADHESION PROTEIN NCAM MODULE 2; CELL ADHESION, GLYCOPROTEIN, HEPARIN-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										BINDING, GPI-ANCHOR, 2 NEURAL ADHESION MOLECULE, IMMUNOGLOBULIN FOLD, HOMOPHILIC 3 BINDING, CELL ADHESION PROTEIN
329	1alh	A	150	220	1.7e-30	-0.17	0.80		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
329	1alh	A	168	248	6.6e-39	-0.12	0.99		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
329	1alh	A	168	250	2e-39			70.01	QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
329	1alh	A	196	277	2e-39	-0.00	0.69		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
329	1alh	A	224	302	1.6e-38	-0.06	0.96		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
329	1alh	A	252	361	1.6e-40	-0.16	0.23		QGSZ ZINC FINGER PEPTIDE; CHAIN: A;	COMPLEX (ZINC FINGER/DNA) COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	(ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
329	1alh	A	310	389	9.9e-42	0.25	0.96		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
329	1alh	A	336	416	6.6e-44	0.02	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
329	1alh	A	364	444	3e-40	-0.18	0.92		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
329	1alh	A	392	473	1.6e-38	-0.15	0.76		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
329	1bbo		153	196	6.6e-17	-0.59	0.03		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11 IBBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) IBBO 4	
329	1bbo		366	420	1.3e-24	-0.40	0.37		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
329	1bbo		394	448	3.3e-23	-0.78	0.10		IBBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) IBBO 4	
329	1mey	G	193	220	6.6e-14	0.36	1.00		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11 IBBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) IBBO 4	
329	1mey	G	389	416	3.3e-13	0.20	0.66		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
329	1mey	G	473	500	3.3e-12	-0.37	0.77		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
329	1mey	G	502	528	1.6e-12	0.39	0.88		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
329	1ubd	C	150	248	3.3e-40	0.04	0.94		YY1; CHAIN: C; ADENO-	COMPLEX (TRANSCRIPTION)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
329	1ubd	C	172	276	3.3e-49	-0.09	1.00		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
329	1ubd	C	193	305	9.9e-52	-0.16	1.00		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
329	1ubd	C	195	305	9.9e-52			83.88	YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
329	1ubd	C	277	416	6.6e-53	-0.17	0.40		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
329	1ubd	C	361	472	6.6e-51	-0.26	0.65		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
329	1ubd	C	389	501	9.9e-52	-0.50	0.89		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
329	1ubd	C	418	528	3.3e-53	-0.06	1.00		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
329	2dtp	A	164	220	9.9e-26	-0.17	0.35		COMPLEX (TRANSCRIPTION REGULATION/DNA) N TRAMTRACK PROTEIN (TWO ZINC-FINGER PEPTIDE) COMPLEXED WITH 2DRP 3 DNA 2DRP 4	COMPLEX (TRANSCRIPTION REGULATION/DNA)
329	2dtp	A	248	302	9.9e-28	0.35	0.49		COMPLEX (TRANSCRIPTION REGULATION/DNA)	COMPLEX (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
329	2dtp	A	276	360	1.6e-27	0.17	0.05		N REGULATION/DNA) TRAMTRACK PROTEIN (TWO ZINC-FINGER PEPTIDE) COMPLEXED WITH 2DRP 3 DNA 2DRP 4	
329	2dtp	A	332	388	1.6e-29	0.48	0.51		COMPLEX(TRANSCRIPTIO N REGULATION/DNA) TRAMTRACK PROTEIN (TWO ZINC-FINGER PEPTIDE) COMPLEXED WITH 2DRP 3 DNA 2DRP 4	
332	1a4y	A	57	374	6.6e-32	0.20	0.98		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPITOPE MAPPING, LEUCINE-RICH 3 REPEATS
332	1a9n	A	136	298	6.6e-16	0.23	-0.01		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP RIBONUCLEOPROTEIN
332	1a9n	A	249	374	1.6e-16	0.05	0.96		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP RIBONUCLEOPROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
332	1a9n	A	60	183	9.9e-22	-0.20	0.16		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
332	1a9n	A	64	207	6.6e-21	0.51	0.51		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
332	1a9n	A	86	242	6.6e-24	0.09	0.90		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
332	1a9n	C	111	309	3.3e-17	0.15	-0.02		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
332	1a9n	C	249	374	1.6e-16	-0.00	0.78		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
332	1cvs	C	404	519	6.6e-19	-0.02	0.24		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
332	1cvs	D	404	519	9.9e-18	0.04	0.18		FIBROBLAST GROWTH	GROWTH FACTOR/GROWTH

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
332	1evt	C	404	519	6.6e-19	0.13	0.58		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
332	1fhg	A	409	501	1.6e-19	0.55	0.95		TELOKIN; CHAIN: A	CONTRACTILE PROTEIN IMMUNOGLOBULIN FOLD, BETA BARREL
332	1nct		410	501	3.3e-21	0.29	0.59		TITIN; CHAIN: NULL;	MUSCLE PROTEIN CONNECTIN, NEXTMS; CELL ADHESION, GLYCOPROTEIN, TRANSMEMBRANE, REPEAT, BRAIN, 2 IMMUNOGLOBULIN FOLD, ALTERNATIVE SPLICING, SIGNAL, 3 MUSCLE PROTEIN
332	1tnm		415	501	9.9e-21	0.53	0.80		MUSCLE PROTEIN TITIN MODULE M5 (CONNECTIN) 1TNM 3 (NMR, MINIMIZED AVERAGE STRUCTURE) 1TNM 4 1TNM 58	
332	1wit		409	500	2.6e-20	0.53	0.18		TWITCHIN 18TH IGSF MODULE; CHAIN: NULL;	MUSCLE PROTEIN IMMUNOGLOBULIN SUPERFAMILY, I SET, MUSCLE PROTEIN
332	1www	X	416	501	3.3e-18	0.05	-0.12		NERVE GROWTH FACTOR; CHAIN: V, W;	NERVE GROWTH FACTOR/TRKA COMPLEX

Table 5

SEQ ID No:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									TRKA RECEPTOR; CHAIN: X, Y;	BETA-NGF; COMPLEX, TRKA RECEPTOR, NERVE GROWTH FACTOR, CYSTEINE KNOT, 2 IMMUNOGLOBULIN LIKE DOMAIN, NERVE GROWTH FACTOR/TRKA COMPLEX ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGEN IN INHIBITOR ACETYLATION, LEUCINE- RICH REPEATS
332	2bnh		60	361	9.9e-44	0.31	0.93		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	
332	3ncm	A	410	500	2.3e-20	0.33	0.80		NEURAL CELL ADHESION MOLECULE, LARGE ISOFORM; CHAIN: A;	CELL ADHESION PROTEIN NCAM MODULE 2; CELL ADHESION, GLYCOPROTEIN, HEPARIN- BINDING, GPI-ANCHOR, 2 NEURAL ADHESION MOLECULE, IMMUNOGLOBULIN FOLD, HOMOPHILIC 3 BINDING, CELL ADHESION PROTEIN
333	1a4i	A	403	511	0.00023	0.33	0.98		METHYLENE TETRAHYDR OFOLATE DEHYDROGENASE / CHAIN: A, B;	OXIDOREDUCTASE METHYLENE THF DEHYDROGENASE / METHENYL THF THF, BIFUNCTIONAL, DEHYDROGENASE, CYCLOHYDROLASE, FOLATE, 2 OXIDOREDUCTASE HEADER
333	1b0a	A	412	511	0.00066	0.22	0.63		FOLD BIFUNCTIONAL PROTEIN; CHAIN: A;	OXIDOREDUCTASE, HYDRO LASE FOLATE, DEHYDROGENASE, CYCLOHYDROLASE, BIFUNCTIONAL, 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
333	1ch6	A	420	517	0.002	0.26	0.24		GLUTAMATE DEHYDROGENASE; CHAIN: A, B, C, D, E, F;	CHANNELING, OXIDOREDUCTASE, HYDROLASE
333	2dlc	A	400	594	1.3e-30	0.01	0.75		D-LACTATE DEHYDROGENASE; 2DLc 5 CHAIN: A, B; 2DLc 6	OXIDOREDUCTASE (CHOH(D)-NAD+(A)) R-LACTATE DEHYDROGENASE; 2DLc 7
338	1alh	A	109	180	2e-16	0.17	-0.01		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
338	1ubd	C	104	180	1.3e-16	-0.14	0.01		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
339	1quu	A	999	1279	9.9e-15	-0.27	0.00		HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
341	1got	B	2	147	6.6e-14	0.47	0.27		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA1

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									GAMMA; CHAIN: G;	SUBUNIT; GAMMA, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
342	1bi8	A	83	204	0.00099	-0.23	0.03		CYCLIN-DEPENDENT KINASE 6; CHAIN: A, C; CYCLIN-DEPENDENT KINASE INHIBITOR; CHAIN: B, D;	COMPLEX (KINASE/INHIBITOR) CDK6; P19INK4D; CYCLIN DEPENDENT KINASE, CYCLIN DEPENDENT KINASE INHIBITORY 2 PROTEIN, CDK, INK4, CELL CYCLE, COMPLEX (KINASE/INHIBITOR) HEADER HELIX
342	1blx	A	83	200	0.0023	-0.01	0.24		CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
342	1cki	A	36	157	0.0066	-0.33	0.17		CASEIN KINASE I DELTA; ICKI 6 CHAIN: A, B; ICKI 7	PHOSPHOTRANSFERASE PROTEIN KINASE ICKI 18
344	1a4f	A	1	141	2e-47	0.65	1.00		HEMOGLOBIN; CHAIN: A, B	OXYGEN TRANSPORT, OXYGEN TRANSPORT, HEME, RESPIRATORY PROTEIN, ERYTHROCYTE
344	1a4f	A	1	141	2e-47			135.78	HEMOGLOBIN; CHAIN: A, B	OXYGEN TRANSPORT, OXYGEN TRANSPORT, HEME, RESPIRATORY PROTEIN, ERYTHROCYTE
344	1a4f	B	1	141	3.3e-38			102.38	HEMOGLOBIN; CHAIN: A, B	OXYGEN TRANSPORT, OXYGEN TRANSPORT, HEME, RESPIRATORY

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
344	1a4f	B	2	140	3.3e-38	0.85	1.00		HEMOGLOBIN; CHAIN: A, B	PROTEIN, ERYTHROCYTE OXYGEN TRANSPORT, HEME, RESPIRATORY
344	1a9w	E	1	140	1.3e-38			106.12	HEMOGLOBIN; CHAIN: A, E, C, F;	PROTEIN, ERYTHROCYTE OXYGEN TRANSPORT
344	1a9w	E	2	140	1.3e-38	0.79	1.00		HEMOGLOBIN; CHAIN: A, E, C, F;	OXYGEN TRANSPORT
344	1bab	A	1	141	9.9e-46			135.21	OXYGEN TRANSPORT HEMOGLOBIN THIONVILLE ALPHA CHAIN MUTANT WITH VAL 1 IBAB 3 REPLACED BY GLU AND AN ACETYLATED MET BOUND TO THE IBAB 4 AMINO TERMINUS IBAB 5	OXYGEN TRANSPORT
344	1bab	A	2	141	9.9e-46	0.78	1.00		OXYGEN TRANSPORT HEMOGLOBIN THIONVILLE ALPHA CHAIN MUTANT WITH VAL 1 IBAB 3 REPLACED BY GLU AND AN ACETYLATED MET BOUND TO THE IBAB 4 AMINO TERMINUS IBAB 5	
344	1c7c	A	1	141	3e-46	0.89	1.00		DEOXYHEMOGLOBIN (ALPHA CHAIN); CHAIN: A; DEOXYHEMOGLOBIN (BETA CHAIN); CHAIN: B, D;	OXYGEN STORAGE/TRANSPORT HEME, OXYGEN DELIVERY VEHICLE, BLOOD SUBSTITUTE
344	1ch4	A	1	141	3.3e-42			105.91	MODULE-SUBSTITUTED CHIMERA HEMOGLOBIN BETA-ALPHA; CHAIN: A, B, C, D;	OXYGEN TRANSPORT, CHIMERA PROTEIN, RESPIRATORY PROTEIN, HEME
344	1ch4	A	2	141	3.3e-42	0.72	1.00		MODULE-SUBSTITUTED CHIMERA HEMOGLOBIN	OXYGEN TRANSPORT OXYGEN TRANSPORT,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
344	1hbb	A	1	141	9.9e-45			120.38	OXYGEN CARRIER HEMOGLOBIN (DEOXY) 1HBB 3	CHIMERA PROTEIN, RESPIRATORY PROTEIN, HEME
344	1hbb	A	2	141	9.9e-45	0.61	1.00		OXYGEN CARRIER HEMOGLOBIN (DEOXY) 1HBB 3	
344	1hbr	A	1	141	9.9e-45	0.93	1.00		HEMOGLOBIN D; CHAIN: A, C; HEMOGLOBIN D; CHAIN: B, D;	OXYGEN STORAGE/TRANSPORT HB D; HB D HEMOGLOBIN D (R- STATE) 1, HEMOGLOBIN, AVIAN, HIGH 2 COOPERATIVITY, OXYGEN TRANSPORT
344	1hbr	A	1	141	9.9e-45			148.19	HEMOGLOBIN D; CHAIN: A, C; HEMOGLOBIN D; CHAIN: B, D;	OXYGEN STORAGE/TRANSPORT HB D; HB D HEMOGLOBIN D (R- STATE) 1, HEMOGLOBIN, AVIAN, HIGH 2 COOPERATIVITY, OXYGEN TRANSPORT
344	1hbr	B	1	138	6.6e-38			96.34	HEMOGLOBIN D; CHAIN: A, C; HEMOGLOBIN D; CHAIN: B, D;	OXYGEN STORAGE/TRANSPORT HB D; HB D HEMOGLOBIN D (R- STATE) 1, HEMOGLOBIN, AVIAN, HIGH 2 COOPERATIVITY, OXYGEN TRANSPORT
344	1hda	A	1	141	3.3e-47	0.49	1.00		OXYGEN TRANSPORT HEMOGLOBIN (DEOXY) 1HDA 3	
344	1hda	A	1	141	3.3e-47			136.61	OXYGEN TRANSPORT HEMOGLOBIN (DEOXY) 1HDA 3	
344	1qpw	A	1	141	3.3e-47	0.78	1.00		PORCINE HEMOGLOBIN (ALPHA SUBUNIT); CHAIN: A, C; PORCINE	OXYGEN TRANSPORT X- RAY STUDY, PORCINE HEMOGLOBIN, ARTIFICIAL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									HEMOGLOBIN (BETA SUBUNIT); CHAIN: B, D	HUMAN BLOOD, 2 OXYGEN TRANSPORT
344	1qpw	A	1	141	3.3e-47			133.84	PORCINE HEMOGLOBIN (ALPHA SUBUNIT); CHAIN: A, C; PORCINE HEMOGLOBIN (BETA SUBUNIT); CHAIN: B, D	OXYGEN TRANSPORT X-RAY STUDY, PORCINE HEMOGLOBIN, ARTIFICIAL HUMAN BLOOD, 2 OXYGEN TRANSPORT
345	1a4f	A	1	179	3.3e-42	0.03	0.28		HEMOGLOBIN; CHAIN: A, B	OXYGEN TRANSPORT OXYGEN TRANSPORT, HEME, RESPIRATORY PROTEIN, ERYTHROCYTE
345	1a4f	A	40	179	3.3e-42			106.59	HEMOGLOBIN; CHAIN: A, B	OXYGEN TRANSPORT OXYGEN TRANSPORT, HEME, RESPIRATORY PROTEIN, ERYTHROCYTE
345	1a4f	B	37	179	2e-32			87.72	HEMOGLOBIN; CHAIN: A, B	OXYGEN TRANSPORT OXYGEN TRANSPORT, HEME, RESPIRATORY PROTEIN, ERYTHROCYTE
345	1a6m		1	179	1.6e-38	0.49	0.49		MYOGLOBIN; CHAIN: NULL;	OXYGEN TRANSPORT HEME PROTEIN, MODEL COMPOUNDS, OXYGEN STORAGE, LIGAND 2 BINDING GEOMETRY, CONFORMATIONAL SUBSTATES, OXYGEN 3 TRANSPORT
345	1a9w	B	37	178	3e-33			93.82	HEMOGLOBIN; CHAIN: A, E, C, F;	OXYGEN TRANSPORT OXYGEN TRANSPORT
345	1bab	A	2	179	2e-40	0.38	1.00		OXYGEN TRANSPORT HEMOGLOBIN THIONVILLE ALPHA CHAIN MUTANT WITH VAL 1 IBAB 3 REPLACED BY GLU AND AN ACETYLATED MET BOUND TO THE IBAB 4 AMINO TERMINUS IBAB 5	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
345	1bab	A	37	179	2e-40			113.33	OXYGEN TRANSPORT HEMOGLOBIN THIONVILLE ALPHA CHAIN MUTANT WITH VAL 1 IBAB 3 REPLACED BY GLU AND AN ACETYLATED MET BOUND TO THE IBAB 4 AMINO TERMINUS IBAB 5	
345	1bab	B	43	179	3.3e-35			86.57	OXYGEN TRANSPORT HEMOGLOBIN THIONVILLE ALPHA CHAIN MUTANT WITH VAL 1 IBAB 3 REPLACED BY GLU AND AN ACETYLATED MET BOUND TO THE IBAB 4 AMINO TERMINUS IBAB 5	
345	1c7c	A	1	179	6.6e-41	0.23	0.25		DEOXYHEMOGLOBIN (ALPHA CHAIN); CHAIN: A; DEOXYHEMOGLOBIN (BETA CHAIN); CHAIN: B, D;	OXYGEN STORAGE/TRANSPORT HEME, OXYGEN DELIVERY VEHICLE, BLOOD SUBSTITUTE
345	1ch4	A	2	179	1.6e-36	0.13	0.99		MODULE-SUBSTITUTED CHIMERA HEMOGLOBIN BETA-ALPHA; CHAIN: A, B, C, D;	OXYGEN TRANSPORT OXYGEN TRANSPORT, CHIMERA PROTEIN, RESPIRATORY PROTEIN, HEME
345	1ch4	A	37	179	1.6e-36			94.07	MODULE-SUBSTITUTED CHIMERA HEMOGLOBIN BETA-ALPHA; CHAIN: A, B, C, D;	OXYGEN TRANSPORT OXYGEN TRANSPORT, CHIMERA PROTEIN, RESPIRATORY PROTEIN, HEME
345	1emy		2	179	6.6e-38	0.44	0.95		OXYGEN TRANSPORT MYOGLOBIN COMPLEXED WITH CYANIDE 1EMY 3 1EMY 107 HEME PROTEIN, GLOBIN FOLD 1EMY 5	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
345	1fth	G	39	179	1.3e-33			100.14	OXYGEN TRANSPORT HEMOGLOBIN (DEOXY, HUMAN FETAL F= β =) 1FDHG 1, 1FDHH 2	
345	1lhb	A	2	179	3.3e-39	0.16	0.36		OXYGEN CARRIER HEMOGLOBIN (DEOXY) 1HBB 3	
345	1lhb	A	39	179	3.3e-39			101.11	OXYGEN CARRIER HEMOGLOBIN (DEOXY) 1HBB 3	
345	1lhb	B	38	179	6.6e-36			76.75	OXYGEN CARRIER HEMOGLOBIN (DEOXY) 1HBB 3	
345	1lhb	A	1	179	3e-39	0.75	1.00		HEMOGLOBIN D; CHAIN: A, C; HEMOGLOBIN D; CHAIN: B, D;	OXYGEN - STORAGE/TRANSPORT HB D; HB D HEMOGLOBIN D (R-STATE) 1, HEMOGLOBIN, AVIAN, HIGH 2 COOPERATIVITY, OXYGEN TRANSPORT
345	1lhb	A	37	179	3e-39			117.48	HEMOGLOBIN D; CHAIN: A, C; HEMOGLOBIN D; CHAIN: B, D;	OXYGEN STORAGE/TRANSPORT HB D; HB D HEMOGLOBIN D (R-STATE) 1, HEMOGLOBIN, AVIAN, HIGH 2 COOPERATIVITY, OXYGEN TRANSPORT
345	1lhb	B	37	176	3.3e-32			81.42	HEMOGLOBIN D; CHAIN: A, C; HEMOGLOBIN D; CHAIN: B, D;	OXYGEN STORAGE/TRANSPORT HB D; HB D HEMOGLOBIN D (R-STATE) 1, HEMOGLOBIN, AVIAN, HIGH 2 COOPERATIVITY, OXYGEN TRANSPORT
345	1hda	A	1	179	1.3e-41	0.49	1.00		OXYGEN TRANSPORT HEMOGLOBIN (DEOXY) 1HDA 3	
345	1hda	A	40	179	1.3e-41			105.99	OXYGEN TRANSPORT HEMOGLOBIN (DEOXY)	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
345	1hda	B	40	179	3.3e-36			82.30	IHDA 3 OXYGEN TRANSPORT HEMOGLOBIN (DEOXY)	
345	1qpw	A	1	179	1.6e-41	0.46	1.00		PORCINE HEMOGLOBIN (ALPHA SUBUNIT); CHAIN: A, C; PORCINE HEMOGLOBIN (BETA SUBUNIT); CHAIN: B, D	OXYGEN TRANSPORT X-RAY STUDY, PORCINE HEMOGLOBIN, ARTIFICIAL HUMAN BLOOD, 2 OXYGEN TRANSPORT
345	1qpw	A	40	179	1.6e-41			110.57	PORCINE HEMOGLOBIN (ALPHA SUBUNIT); CHAIN: A, C; PORCINE HEMOGLOBIN (BETA SUBUNIT); CHAIN: B, D	OXYGEN TRANSPORT X-RAY STUDY, PORCINE HEMOGLOBIN, ARTIFICIAL HUMAN BLOOD, 2 OXYGEN TRANSPORT
345	1qpw	B	37	179	2e-35			88.40	PORCINE HEMOGLOBIN (ALPHA SUBUNIT); CHAIN: A, C; PORCINE HEMOGLOBIN (BETA SUBUNIT); CHAIN: B, D	OXYGEN TRANSPORT X-RAY STUDY, PORCINE HEMOGLOBIN, ARTIFICIAL HUMAN BLOOD, 2 OXYGEN TRANSPORT
345	1tln	B	38	179	3e-36			80.38	HEMOGLOBIN; CHAIN: A; HEMOGLOBIN; CHAIN: B;	OXYGEN TRANSPORT HBTNCO; HBTNCO; OXYGEN TRANSPORT, HEMOGLOBIN
348	1a0a	A	83	136	0.00033	-0.25	0.55		PHOSPHATE SYSTEM POSITIVE REGULATORY PROTEIN CHAIN: A, B; DNA; CHAIN: C, D;	COMPLEX (TRANSCRIPTION FACTOR/DNA) BHLH; UASP2(17); TRANSCRIPTION FACTOR, BASIC HELIX LOOP HELIX, 2 COMPLEX (TRANSCRIPTION FACTOR/DNA)
348	1am9	A	83	149	1.3e-07	-0.46	0.06		STEROL REGULATORY ELEMENT BINDING PROTEIN 1A; CHAIN: A, B, C, D; DNA; CHAIN: E, F, G, H;	COMPLEX (TRANSCRIPTION REGULATION/DNA) SREBP-1A; STEROL REGULATORY ELEMENT BINDING PROTEIN, 2 BASIC-HELIX-LOOP-HELIX-LEUCINE ZIPPER, SREBP, TRANSCRIPTION 3 FACTOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
348	1am9	B	83	153	3.3e-07	-0.55	0.29		STEROL REGULATORY ELEMENT BINDING PROTEIN 1A; CHAIN: A, B, C, D; DNA; CHAIN: E, F, G, H;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA) SREBP-1A; STEROL REGULATORY ELEMENT BINDING PROTEIN, 2 BASIC-HELIX-LOOP-HELIX-LEUCINE ZIPPER, SREBP, TRANSCRIPTION 3 FACTOR, COMPLEX (TRANSCRIPTION REGULATION/DNA)
352	1chc		193	223	0.0023	-0.78	0.00		VIRUS EQUINE HERPES VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	
352	1chc		29	80	0.00099	0.03	0.28		VIRUS EQUINE HERPES VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	
352	1g25	A	191	224	2.3e-05	-0.76	0.41		CDK-ACTIVATING KINASE ASSEMBLY FACTOR MAT1; CHAIN: A;	METAL BINDING PROTEIN RING FINGER PROTEIN MAT1; RING FINGER (C3HC4)
353	1chc		158	188	0.0023	-0.78	0.00		VIRUS EQUINE HERPES VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	
353	1dvp	A	153	187	0.0097	-0.44	0.41		HEPATOCYTE GROWTH FACTOR-REGULATED TYROSINE CHAIN: A;	TRANSFERASE HRS; HRS, VHS, FYVE, ZINC FINGER, SUPERHELIX
353	1g25	A	156	189	2.3e-05	-0.76	0.41		CDK-ACTIVATING KINASE ASSEMBLY FACTOR MAT1; CHAIN: A;	METAL BINDING PROTEIN RING FINGER PROTEIN MAT1; RING FINGER (C3HC4)
354	1cfe		53	212	3.3e-41			84.39	PATHOGENESIS-RELATED	PATHOGENESIS-RELATED

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									PROTEIN P14A; CHAIN: NULL;	PROTEIN PATHOGENESIS-RELATED LEAF PROTEIN 6, ETHYLENE PATHOGENESIS-RELATED PROTEIN, PR-1 PROTEINS, 2 PLANT DEFENSE
354	1cfe		53	212	5.1e-35	0.30	0.99		PATHOGENESIS-RELATED PROTEIN P14A; CHAIN: NULL;	PATHOGENESIS-RELATED PROTEIN PATHOGENESIS-RELATED LEAF PROTEIN 6, ETHYLENE PATHOGENESIS-RELATED PROTEIN, PR-1 PROTEINS, 2 PLANT DEFENSE
354	1cfe		56	212	3.3e-41	0.40	0.95		PATHOGENESIS-RELATED PROTEIN P14A; CHAIN: NULL;	PATHOGENESIS-RELATED PROTEIN PATHOGENESIS-RELATED LEAF PROTEIN 6, ETHYLENE PATHOGENESIS-RELATED PROTEIN, PR-1 PROTEINS, 2 PLANT DEFENSE
354	1qnx	A	24	212	8.5e-39	0.23	1.00		VES V 5; CHAIN: A;	ALLERGEN ANTIGEN 5; ANTIGEN 5, ALLERGEN, VESPID VENOM
354	1qnx	A	47	213	3.3e-42	0.31	1.00		VES V 5; CHAIN: A;	ALLERGEN ANTIGEN 5; ANTIGEN 5, ALLERGEN, VESPID VENOM
357	1d6j	A	17	58	3.3e-05	0.03	0.25		ADENOSINE-5'PHOSPHOSULFATE KINASE; CHAIN: A, B;	TRANSFERASE APS KINASE; APS KINASE, ADENYLYLSULFATE KINASE, SULFATE, NUCLEOTIDE 2 KINASE, TRANSFERASE
359	3znf		143	172	0.00051	0.26	0.13		ZINC FINGER/DNA\$ BINDING DOMAIN ZINC FINGER (NMR\$) 3ZNF 3	
360	1191		81	154	1.6e-15	0.02	0.21		HYDROLASE(O-	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
360	1191		81	203	5.1e-11	-0.12	0.11		GLYCOSYL) LYSOZYME (E.C.3.2.1.17) MUTANT WITH CYS 54 REPLACED BY THR, 119L 3 CYS 97 REPLACED BY ALA, ALA 134 REPLACED BY SER (C54T,C97A, 119L 4 A134S) 119L 5	
									HYDROLASE(O-GLYCOSYL) LYSOZYME (E.C.3.2.1.17) MUTANT WITH CYS 54 REPLACED BY THR, 119L 3 CYS 97 REPLACED BY ALA, ALA 134 REPLACED BY SER (C54T,C97A, 119L 4 A134S) 119L 5	
360	1741	A	81	203	1.4e-11	-0.09	0.03		HYDROLASE (O-GLYCOSYL) LYSOZYME (E.C.3.2.1.17) MUTANT WITH THR 34 REPLACED BY ALA, 174L 3 LYS 35 REPLACED BY ALA, SER 36 REPLACED BY ALA, PRO 37 174L 4 REPLACED BY ALA, SER 38 REPLACED BY ASP, ASN 40 REPLACED BY 174L 5 ALA, SER 44 REPLACED BY ALA, GLU 45 REPLACED BY ALA, ASP 47 174L 6 REPLACED BY ALA, LYS 48 REPLACED BY ALA, CYS 54 REPLACED BY 174L 7 THR, CYS 97 REPLACED BY ALA	
									CT34A,K35A,S36A,P37A,S38 D,N40A, 174L 8	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
360	176l	A	80	154	1.6e-15	-0.04	0.40		S44A,E45A,D47A,K48A,C54 T,C97A) 176L 9 HYDROLASE (O-GLYCOSYL) LYSOZYME (E.C.3.2.1.17) MUTANT WITH LEU 32 REPLACED BY THR, 176L 3 THR 34 REPLACED BY LYS, LYS 35 REPLACED BY VAL, SER 36 176L 4 REPLACED BY ASP, PRO 37 REPLACED BY GLY, SER 38 REPLACED BY 176L 5 ASN, LEU 39 REPLACED BY SER, CYS 54 REPLACED BY THR, AND 176L 6 CYS 97 REPLACED BY ALA (L32T,T34K,K35V,S36D,P37 G,S38N, 176L 7 L39S,C54T,C97A) 176L 8	
360	176l	A	81	203	1.7e-12	0.10	0.03		HYDROLASE (O-GLYCOSYL) LYSOZYME (E.C.3.2.1.17) MUTANT WITH LEU 32 REPLACED BY THR, 176L 3 THR 34 REPLACED BY LYS, LYS 35 REPLACED BY VAL, SER 36 176L 4 REPLACED BY ASP, PRO 37 REPLACED BY GLY, SER 38 REPLACED BY 176L 5 ASN, LEU 39 REPLACED BY SER, CYS 54 REPLACED BY THR, AND 176L 6 CYS 97 REPLACED BY ALA (L32T,T34K,K35V,S36D,P37 G,S38N, 176L 7	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
360	189l		80	154	1.3e-15	-0.01	0.35		L39S,C54T,C97A) 176L 8 HYDROLASE (O-GLYCOSYL) LYSOZYME (E.C.3.2.1.17) MUTANT WITH ILE 3 REPLACED BY LEU, 189L 3 SER 38 REPLACED BY ASP, ALA 41 REPLACED BY VAL, ALA 82 189L 4 PRO, ASN 116 REPLACED BY ASP, VAL 131 REPLACED BY ALA, 189L 5 AND ASN 144 REPLACED BY ASP SUBSTITUTIONS (13L,S38D,A41V, 189L 6 A82P,N116D,V131A,N144D) 189L 7	
360	191l		81	154	1.3e-16	0.11	0.51		LYSOZYME; 191L 4 CHAIN: NULL; 191L 5	HYDROLASE (O-GLYCOSYL)
360	191l		81	203	1.7e-11	0.13	0.21		LYSOZYME; 191L 4 CHAIN: NULL; 191L 5	HYDROLASE (O-GLYCOSYL)
360	192l		81	203	3.4e-11	-0.28	0.03		LYSOZYME; 192L 4 CHAIN: NULL; 192L 5	HYDROLASE (O-GLYCOSYL)
360	205l		81	203	1e-11	0.15	0.43		HYDROLASE(O-GLYCOSYL) LYSOZYME (E.C.3.2.1.17) INSERTION MUTANT WITH ALA ALA 205L 3 INSERTED AFTER SER 44, CYS 54 REPLACED BY THR, CYS 97 205L 4 REPLACED BY ALA (INS(S44-AAA),C54T,C97A) 205L 5	
361	1b7y	B	2	416	5.1e-52	-0.38	0.00		PHENYLALANYL-TRNA SYNTHETASE; CHAIN: A; PHENYLALANYL-TRNA SYNTHETASE; CHAIN: B;	LIGASE PHERS; PHERS; ENZYME, TRNA SYNTHETASE, ALPHA/BETA HOMODIMER

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
364	1bor		282	338	2.3e-12	-0.89	0.09		TRANSCRIPTION FACTOR PML; CHAIN: NULL;	TRANSCRIPTION REGULATION PROTO-ONCOGENE, NUCLEAR BODIES (PODS), LEUKEMIA, 2 TRANSCRIPTION REGULATION
364	1bor		283	332	1.7e-05	-0.59	0.04		TRANSCRIPTION FACTOR PML; CHAIN: NULL;	TRANSCRIPTION REGULATION PROTO-ONCOGENE, NUCLEAR BODIES (PODS), LEUKEMIA, 2 TRANSCRIPTION REGULATION
364	1chc		288	342	6.6e-16	0.19	0.62		VIRUS EQUINE HERPES VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	
364	1chc		289	333	1.7e-14	-0.01	1.00		VIRUS EQUINE HERPES VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	
364	1g25	A	288	338	6.6e-14	-0.27	0.07		CDK-ACTIVATING KINASE ASSEMBLY FACTOR MAT1; CHAIN: A;	METAL BINDING PROTEIN RING FINGER PROTEIN MAT1; RING FINGER (C3HC4)
364	1rmd		284	332	1.4e-07	-0.08	0.70		RAG1; CHAIN: NULL;	DNA-BINDING PROTEIN V(D)J RECOMBINATION ACTIVATING PROTEIN 1; RAG1, V(D)J RECOMBINATION, ANTIBODY, MAD, RING FINGER, 2 ZINC BINUCLEAR CLUSTER, ZINC FINGER, DNA-BINDING PROTEIN
365	1chc		109	155	6.8e-05	-0.07	0.84		VIRUS EQUINE HERPES VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	
365	1chc		109	155	9.9e-06	-0.07	0.84		VIRUS EQUINE HERPES	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
365	1g25	A	117	152	3.3e-06	-0.14	0.30		VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	METAL BINDING PROTEIN RING FINGER PROTEIN MAT1; RING FINGER (C3HC4)
365	1rmd		102	186	0.001	-0.76	0.09		CDK-ACTIVATING KINASE ASSEMBLY FACTOR MAT1; CHAIN: A; RAG1; CHAIN: NULL;	DNA-BINDING PROTEIN V(D)J RECOMBINATION ACTIVATING PROTEIN 1; RAG1, V(D)J RECOMBINATION, ANTIBODY, MAD, RING FINGER, 2 ZINC BINUCLEAR CLUSTER, ZINC FINGER, DNA-BINDING PROTEIN
367	1elw	A	3	39	5.1e-07	-0.28	0.68		TPR1-DOMAIN OF HOP; CHAIN: A, B; HSC70-PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
369	1aip	A	20	220	1.7e-38	0.07	-0.18		ELONGATION FACTOR TU; CHAIN: A, B, E, F; ELONGATION FACTOR TS; CHAIN: C, D, G, H;	COMPLEX OF TWO ELONGATION FACTORS EF-TU; EF-TS; ELONGATION FACTOR, NUCLEOTIDE EXCHANGE, GTP-BINDING, 2 COMPLEX OF TWO ELONGATION FACTORS
369	1d2e	A	20	219	1.4e-33	0.40	-0.15		ELONGATION FACTOR TU (EF-TU); CHAIN: A, B, C, D	RNA BINDING PROTEIN G-PROTEIN, BETA-BARREL
369	1dar		11	155	1.4e-12	0.18	-0.07		ELONGATION FACTOR G; CHAIN: NULL;	TRANSLATIONAL GTPASE EF-G RIBOSOMAL TRANSLOCASE, TRANSLATIONAL GTPASE
369	1efc	A	20	219	3.4e-41	0.12	-0.15		ELONGATION FACTOR; CHAIN: A, B;	RNA BINDING PROTEIN EFTU; TRANSPORT AND PROTECTION PROTEIN, RNA

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										BINDING PROTEIN
369	1ega	A	25	225	3.4e-35	0.12	-0.12		GTP-BINDING PROTEIN ERA; CHAIN: A, B;	HYDROLASE ERA, GTPASE, RNA-BINDING, RAS-LIKE, HYDROLASE
369	1etu		16	208	1.7e-37	0.21	-0.14		TRANSPORT AND PROTECTION PROTEIN ELONGATION FACTOR TU (DOMAIN I) - *GUANOSINE DIPHOSPHATE IETU 4 COMPLEX IETU 5	
369	1exm	A	16	220	1.7e-43	0.09	-0.12		ELONGATION FACTOR TU (EF-TU); CHAIN: A;	TRANSLATION EF-TU; GTPASE, MOLECULAR SWITCH, TRNA, RIBOSOME, Q-BETA REPLICASE, 2 CHAPERONE, DISULFID ISOMERASE
369	1fnn	A	19	155	1.7e-12	0.17	-0.14		ELONGATION FACTOR G; CHAIN: A;	TRANSLATION EF-G; BENT CONFORMATION, VISIBLE DOMAIN III, MUTATION HIS573ALA
369	1g7s	A	21	213	6.8e-18	0.23	-0.09		TRANSLATION INITIATION FACTOR IF2/EIF5B; CHAIN: A;	TRANSLATIONAL GTPASE
369	2efg	A	19	155	1.7e-12	0.19	-0.14		ELONGATION FACTOR G; CHAIN: A; ELONGATION FACTOR G DOMAIN 3; CHAIN: B;	PROTEIN BINDING EF-G; EF-G ELONGATION FACTOR, TRANSLOCASE, RIBOSOME, ELONGATION, 2 TRANSLATION, PROTEIN SYNT FACTOR, GTPASE, GTP BINDING, 3 GUANOSINE NUCLEOTIDE BINDING, PROTEIN BINDING
370	1alh	A	153	262	9.9e-42	-0.28	0.23		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
370	1alh	A	237	319	3.3e-37			81.16	OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	FINGER, DNA-BINDING PROTEIN
370	1alh	A	237	346	3.3e-37	-0.13	0.17		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
370	1alh	A	349	430	9.9e-36	0.40	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
370	1alh	A	377	457	6.6e-37	0.42	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
370	1alh	A	405	486	3.3e-36	0.30	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
370	1alh	A	433	513	1.6e-36	0.47	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
370	1alh	A	97	177	1.7e-25	0.20	0.13		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
370	1mey	C	124	205	1.7e-44	0.49	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
370	1mey	C	152	233	6.8e-46	0.55	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION; PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
370	1mey	C	180	261	1e-46	0.44	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
370	1mey	C	208	289	1.2e-47	0.38	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
370	1mey	C	236	317	1.2e-48	0.08	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
370	1mey	C	236	318	1.5e-49			107.60	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	STRUCTURE, COMPLEX (ZINC FINGER/DNA)
370	1mey	C	264	345	1.5e-49	0.34	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
370	1mey	C	292	373	1.4e-50	0.36	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
370	1mey	C	320	401	1.7e-50	0.32	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
370	1mey	C	348	429	3.4e-50	0.47	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
370	1mey	C	376	457	3.4e-50	0.68	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER	COMPLEX (ZINC FINGER/DNA) ZINC FINGER

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
370	1mey	C	404	485	5.1e-50	0.40	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
370	1mey	C	432	513	6.8e-50	0.53	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
370	1mey	C	460	517	1.4e-33	0.46	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
370	1mey	C	96	177	5.1e-42	-0.01	0.36		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
370	1mey	G	150	177	3.4e-11	0.54	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
370	1mey	G	150	177	6.6e-15	0.54	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	(ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
370	1tf6	A	122	287	5.1e-37			114.40	TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
370	1tf6	A	125	270	1.4e-35	0.41	1.00		TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
370	1tf6	A	181	326	5.1e-37	0.18	1.00		TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
370	1tf6	A	265	410	1.5e-38	0.03	1.00		TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
370	1tf6	A	349	495	1.7e-37	0.17	0.95		TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	PROTEIN COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
370	1tf6	A	377	515	3.4e-36	0.50	1.00		TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
370	1ubd	C	127	233	3.4e-31	0.42	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
370	1ubd	C	127	261	6.6e-53	0.30	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
370	1ubd	C	160	261	3.4e-32	0.44	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR

Table 5.

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
370	1ubd	C	213	318	3.3e-46	0.15	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
370	1ubd	C	216	317	5.1e-33	0.22	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
370	1ubd	C	244	345	3.4e-34	0.23	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
370	1ubd	C	266	373	1.4e-35	0.50	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
370	1ubd	C	294	402	1.3e-48			96.52	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
370	1ubd	C	297	402	1.3e-48	0.05	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
370	1ubd	C	300	401	5.1e-36	0.20	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
370	1ubd	C	346	458	1.3e-47	0.06	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
370	1ubd	C	356	457	5.1e-35	0.46	1.00		YY1; CHAIN: C; ADENO-	COMPLEX (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
370	1ubd	C	374	513	6.6e-46	0.15	0.99		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
370	1ubd	C	381	485	5.1e-35	0.16	1.00		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
370	1ubd	C	412	513	1.7e-34	0.04	1.00		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
370	1ubd	C	440	517	1.7e-25	0.18	1.00		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
370	2gli	A	124	260	1.4e-31	0.58	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
370	2gli	A	180	319	3.4e-34			99.97	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
370	2gli	A	216	344	3.4e-34	0.29	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
370	2gli	A	384	512	1.5e-34	0.45	0.99		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
370	2gli	A	89	207	1.7e-29	0.09	0.64		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
371	1byu	A	19	186	1.7e-48			55.66	GTP-BINDING PROTEIN RAN; CHAIN: A, B;	TRANSPORT PROTEIN TC4; GTPASE, NUCLEAR TRANSPORT, TRANSPORT PROTEIN
371	1byu	B	20	181	1.7e-48	0.62	1.00		GTP-BINDING PROTEIN RAN; CHAIN: A, B;	TRANSPORT PROTEIN TC4; GTPASE, NUCLEAR TRANSPORT, TRANSPORT PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
371	1cly	A	18	185	1.5e-52			58.13	RAS-RELATED PROTEIN RAP-1A; CHAIN: A; PROTO-ONCOGENE SERINE/THREONINE PROTEIN KINASE CHAIN: B;	SIGNALING PROTEIN GTP-BINDING PROTEINS, PROTEIN-PROTEIN COMPLEX, EFFECTORS
371	1cly	A	19	181	1.5e-52	0.18	1.00		RAS-RELATED PROTEIN RAP-1A; CHAIN: A; PROTO-ONCOGENE SERINE/THREONINE PROTEIN KINASE CHAIN: B;	SIGNALING PROTEIN GTP-BINDING PROTEINS, PROTEIN-PROTEIN COMPLEX, EFFECTORS
371	1ctq	A	18	184	3.4e-55			60.17	TRANSFORMING PROTEIN P21/H-RAS-1; CHAIN: A;	SIGNALING PROTEIN G PROTEIN, GTP HYDROLYSIS, KINETIC CRYSTALLOGRAPHY, 2 SIGNALING PROTEIN
371	1ctq	A	20	183	3.4e-55	0.59	0.98		TRANSFORMING PROTEIN P21/H-RAS-1; CHAIN: A;	SIGNALING PROTEIN G PROTEIN, GTP HYDROLYSIS, KINETIC CRYSTALLOGRAPHY, 2 SIGNALING PROTEIN
371	1cxz	A	16	186	3.4e-37			55.83	HIS-TAGGED TRANSFORMING PROTEIN RHOA(0-181); CHAIN: A; PKN; CHAIN: B;	SIGNALING PROTEIN PROTEIN-PROTEIN COMPLEX, ANTIPARALLEL COILED-COIL
371	1d5c	A	20	181	1.5e-52	0.74	1.00		RAB6 GTPASE; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS G-PROTEIN, GTPASE, RAB6, VESICULAR TRAFFICKING
371	1dar		31	137	3.4e-05	-0.16	0.24		ELONGATION FACTOR G; CHAIN: NULL;	TRANSLATIONAL GTPASE EF-G RIBOSOMAL TRANSLOCASE, TRANSLATIONAL GTPASE
371	1e0s	A	12	182	3.4e-56	0.81	1.00		ADP-RIBOSYLATION FACTOR 6; CHAIN: A;	G PROTEIN G PROTEIN, RAS, ARF, ARF6, MEMBRANE TRAFFIC
371	1ek0	A	22	181	6.8e-52	0.63	1.00		GTP-BINDING PROTEIN YPT51; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS G PROTEIN, VESICULAR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
371	1fmm	A	31	137	1.7e-05	-0.11	0.03		ELONGATION FACTOR G; CHAIN: A;	TRAFFIC, GTP HYDROLYSIS, YPT/RAB 2 PROTEIN, ENDOCYTOSIS, HYDROLASE
371	1hur	A	5	185	1e-64	0.98	1.00		HUMAN ADP-RIBOSYLATION FACTOR 1; 1HUR 5 CHAIN: A, B; 1HUR 7	TRANSLATION EF-G; BENT CONFORMATION, VISIBLE DOMAIN III, MUTATION HIS573ALA
371	1hur	A	6	185	1e-64			138.57	HUMAN ADP-RIBOSYLATION FACTOR 1; 1HUR 5 CHAIN: A, B; 1HUR 7	PROTEIN TRANSPORT GDP-BINDING, MEMBRANE TRAFFICKIN, NON-MYRISTOYLATED 1HUR 16
371	1ibr	A	20	186	6.8e-48			77.80	RAN; CHAIN: A, C; IMPORTIN BETA SUBUNIT; CHAIN: B, D; RAP2A; CHAIN: NULL;	PROTEIN TRANSPORT GDP-BINDING, MEMBRANE TRAFFICKIN, NON-MYRISTOYLATED 1HUR 16
371	1kao		18	184	3.4e-51			58.03		SMALL GTPASE KARYOPHERIN BETA, P95 SMALL GTPASE, NUCLEAR TRANSPORT RECEPTOR
371	1kao		19	181	3.4e-51	0.49	1.00			GTP-BINDING PROTEIN, GTP-BINDING PROTEIN, SMALL G PROTEIN, RAP2, GDP, RAS
371	1trp	C	18	186	6.8e-48			69.18	RAN; CHAIN: A, C; NUCLEAR PORE COMPLEX PROTEIN NUP358; CHAIN: B, D;	GTP-BINDING PROTEIN, GTP-BINDING PROTEIN, SMALL G PROTEIN, RAP2, GDP, RAS
371	1tx4	B	19	184	1e-34			51.89	P50-RHO GAP; CHAIN: A; TRANSFORMING PROTEIN RHOA; CHAIN: B;	COMPLEX (SMALL GTPASE/NUCLEAR PROTEIN) COMPLEX (SMALL GTPASE/NUCLEAR PROTEIN), SMALL GTPASE, 2 NUCLEAR TRANSPORT

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
371	1zbd	A	13	186	1.7e-57	0.74	1.00		RAB-3A; CHAIN: A; RABPHILIN-3A; CHAIN: B;	(GTPASE ACTIVATION/PROTO- ONCOGENE), GTPASE, 2 TRANSITION STATE, GAP COMPLEX (GTP- BINDING/EFFECTOR) RAS- RELATED PROTEIN RAB3A; COMPLEX (GTP- BINDING/EFFECTOR), G PROTEIN, EFFECTOR, RABCDR, 2 SYNAPTIC EXOCYTOSIS, RAB PROTEIN, RAB3A, RABPHILIN COMPLEX (GTP- BINDING/EFFECTOR) RAS- RELATED PROTEIN RAB3A; COMPLEX (GTP- BINDING/EFFECTOR), G PROTEIN, EFFECTOR, RABCDR, 2 SYNAPTIC EXOCYTOSIS, RAB PROTEIN, RAB3A, RABPHILIN COMPLEX (GTP- BINDING/EFFECTOR) RAS- RELATED PROTEIN RAB3A; COMPLEX (GTP- BINDING/EFFECTOR), G PROTEIN, EFFECTOR, RABCDR, 2 SYNAPTIC EXOCYTOSIS, RAB PROTEIN, RAB3A, RABPHILIN PROTEIN BINDING EF-G; EF- G ELONGATION FACTOR, TRANSLOCASE, RIBOSOME, ELONGATION, 2 TRANSLATION, PROTEIN SYNT FACTOR, GTPASE, GTP BINDING, 3 GUANOSINE NUCLEOTIDE BINDING, PROTEIN BINDING HYDROLASE G PROTEIN, VESICULAR TRAFFICKING, GTP HYDROLYSIS, RAB 2 PROTEIN, NEUROTRANSMITTER
371	1zbd	A	18	186	1.7e-57			63.80	RAB-3A; CHAIN: A; RABPHILIN-3A; CHAIN: B;	
371	2efg	A	31	137	3.4e-05	-0.05	0.35		ELONGATION FACTOR G; CHAIN: A; ELONGATION FACTOR G DOMAIN 3; CHAIN: B;	
371	3rab	A	17	184	8.5e-58			71.97	RAB3A; CHAIN: A;	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
371	3rab	A	18	184	8.5e-58	0.72	1.00		RAB3A; CHAIN: A;	RELEASE, HYDROLASE HYDROLASE G PROTEIN, VESICULAR TRAFFICKING, GTP HYDROLYSIS, RAB 2 PROTEIN, NEUROTRANSMITTER RELEASE, HYDROLASE
372	1eur		49	415	1.7e-53			135.41	SIALIDASE; CHAIN: NULL;	HYDROLASE NEURAMINIDASE; NEURAMINIDASE, SIALIDASE, HYDROLASE
372	1eur		58	336	1.7e-53	0.48	0.93		SIALIDASE; CHAIN: NULL;	HYDROLASE NEURAMINIDASE; NEURAMINIDASE, SIALIDASE, HYDROLASE
372	1eut		49	528	0.00017			99.82	SIALIDASE; CHAIN: NULL;	HYDROLASE NEURAMINIDASE; HYDROLASE, GLYCOSIDASE
372	1eut		58	336	1.7e-53	0.46	0.99		SIALIDASE; CHAIN: NULL;	HYDROLASE NEURAMINIDASE; HYDROLASE, GLYCOSIDASE
372	1kit		147	304	0.0001	0.43	0.10		SIALIDASE; CHAIN: NULL;	HYDROLASE NEURAMINIDASE; HYDROLASE, GLYCOSIDASE, SIGNAL, REPEAT, CALCIUM
372	1kit		171	334	3.3e-17	0.13	-0.01		SIALIDASE; CHAIN: NULL;	HYDROLASE NEURAMINIDASE; HYDROLASE, GLYCOSIDASE, SIGNAL, REPEAT, CALCIUM
372	1kit		204	305	3.4e-12	0.39	0.23		SIALIDASE; CHAIN: NULL;	HYDROLASE NEURAMINIDASE; HYDROLASE, GLYCOSIDASE, SIGNAL, REPEAT, CALCIUM

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
372	2sli		168	341	6.8e-32	0.30	0.05		INTRAMOLECULAR TRANS-SIALIDASE; CHAIN: NULL;	HYDROLASE HYDROLASE, INTRAMOLECULAR TRANS-SIALIDASE, NEURAMINIDASE
372	3sil		33	417	8.5e-57			92.46	SIALIDASE; CHAIN: NULL;	GLYCOSIDASE GLYCOSIDASE, HYDROLASE
372	3sil		45	328	8.5e-57	0.38	0.03		SIALIDASE; CHAIN: NULL;	GLYCOSIDASE GLYCOSIDASE, HYDROLASE
373	1alh	A	138	228	3.4e-18	-0.32	0.25		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
373	1alh	A	202	290	3.4e-26			65.40	QGSZ ZINC FINGER PEPTIDE; CHAIN: A; OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
373	1alh	A	232	316	3.4e-26	0.35	0.35		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
373	1bbo		264	315	3.4e-07	-0.20	0.04		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11 IBBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) IBBO 4	
373	1mey	C	136	228	1.7e-36	-0.27	0.18		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
373	1mey	C	201	289	3.4e-41			71.51	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
373	1mey	C	231	316	3.4e-41	0.10	0.27		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
373	1uf3	A	201	292	5.1e-21			69.81	TRANSCRIPTION FACTOR III _A ; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIII _A ; 5S GENE; NMR, TFIII _A , PROTEIN; DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
373	1uf3	A	232	315	5.1e-21	0.15	-0.01		TRANSCRIPTION FACTOR III _A ; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIII _A ; 5S GENE; NMR, TFIII _A , PROTEIN; DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
373	1uf6	A	101	236	3.4e-24	-0.33	0.00		TFIII _A ; CHAIN: A, D; 5S	COMPLEX (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
373	1tf6	A	137	267	1.5e-27	-0.09	0.04		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
373	1tf6	A	165	315	5.1e-38	0.04	0.70		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
373	1tf6	A	168	334	5.1e-38			95.69	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
373	1ubd	C	113	228	3.4e-26	-0.34	0.24		YY1; CHAIN: C; ADENO- ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING- YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA- PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
373	1ubd	C	144	258	6.6e-34	0.13	0.98		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	REGULATION/DNA COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
373	1ubd	C	176	289	1.2e-34			89.17	YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
373	1ubd	C	206	315	1.2e-34	0.10	0.60		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
373	2adr		264	315	1.2e-10	0.03	0.03		ADRI; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADRI, ZINC FINGER, NMR
373	2gli	A	136	289	6.8e-53			157.30	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
373	2gli	A	168	287	6.8e-53	0.30	1.00		ZINC FINGER PROTEIN	COMPLEX (DNA-BINDING PROTEIN/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
373	2gli	A	173	315	3.4e-41	0.27	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
374	1cex		506	639	0.002	-0.15	0.01		CUTINASE; CHAIN: NULL;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
376	1b7f	A	108	172	5.1e-13	0.63	1.00		SXL-LETHAL PROTEIN; CHAIN: A, B; RNA (5'-R(P*Gp*Up*Up*Gp*Up*Up*R(P*Gp*Up*Up*Gp*Up*U)-CHAIN: P, Q); POLYDENYLATE BINDING PROTEIN I; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	RNA-BINDING PROTEIN/RNA TRA PRE-MRNA; SPLICING REGULATION, RNP DOMAIN, RNA COMPLEX GENE REGULATION/RNA POLY(A) BINDING PROTEIN-1, PAPB 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
376	1d8z	A	110	173	6.6e-14	0.17	0.95		HU ANTIGEN C; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
376	1fj7	A	100	173	1.3e-13	0.47	0.71		NUCLEOLIN RBD1; CHAIN: A;	STRUCTURAL PROTEIN C23; RNP, RBD, RRM, RNA BINDING DOMAIN, NUCLEOLUS
376	1fjc	A	100	173	3.3e-14	0.51	0.74		NUCLEOLIN RBD2; CHAIN: A;	STRUCTURAL PROTEIN C23; RNP, RBD, RRM, RNA BINDING DOMAIN, NUCLEOLUS
376	1ha1		105	170	1.7e-18	0.35	0.80		HNRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEOP

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
376	1hd1	A	111	166	8.5e-13	0.61	0.30		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN D0; CHAIN: A;	RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
376	1sxl		105	173	6.6e-13	0.60	0.59		RNA-BINDING PROTEIN (C-SEX-LETHAL PROTEIN (C-TERMINUS, OR SECOND RNA-BINDING DOMAIN 1SXL 3 (RBD-2), RESIDUES 199 - 294 PLUS N-TERMINAL MET) 1SXL 4 (NMR, 17 STRUCTURES) 1SXL 5	RNA-BINDING DOMAIN RNA-BINDING DOMAIN, ALTERNATIVE SPLICING
376	2sxl		108	172	5.1e-13	0.89	1.00		SEX-LETHAL PROTEIN; CHAIN: NULL;	COMPLEX (RIBONUCLEOPROTEIN/DNA) HNRNP A1, UPI; COMPLEX (RIBONUCLEOPROTEIN/DNA), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1
376	2up1	A	104	173	6.8e-20	0.65	0.72		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1; CHAIN: A; 12-NUCLEOTIDE SINGLE-STRANDED TELOMETRIC DNA; CHAIN: B;	RNA-BINDING
377	1b7f	A	110	192	5.1e-22	0.77	1.00		SXL-LETHAL PROTEIN; CHAIN: A, B; RNA (5'-R(P*Gp*Up*Up*Gp*Up*Up*Up*Up*Up*Up*U)-CHAIN: P, Q;	PROTEIN/RNA TRA PRE-MRNA; SPLICING REGULATION, RNP DOMAIN, RNA COMPLEX
377	1d8z	A	107	192	3.4e-22	0.87	1.00		HU ANTIGEN C; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
377	1d8z	A	110	193	9.9e-25	0.35	1.00		HU ANTIGEN C; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
377	1d9a	A	111	193	2.3e-22	0.86	1.00		HU ANTIGEN C; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
377	1f7	A	100	193	9.9e-23	0.34	0.49		NUCLEOLIN RBD1; CHAIN: A;	STRUCTURAL PROTEIN PROTEIN C23; RNP, RBD, RRM, RNA BINDING DOMAIN, NUCLEOLUS
377	1ha1		105	190	1.7e-30	0.95	1.00		HNRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
377	1hd1	A	111	186	1.4e-24	1.14	0.99		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN D0; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
377	1sxl		105	193	6.6e-24	0.96	0.99		RNA-BINDING PROTEIN SEX-LETHAL PROTEIN (C- TERMINUS, OR SECOND RNA-BINDING DOMAIN ISXL 3 (RBD-2), RESIDUES 199 - 294 PLUS N- TERMINAL MET) ISXL 4 (NMR, 17 STRUCTURES) ISXL 5	
377	1sxl		98	191	6.6e-24			51.72	RNA-BINDING PROTEIN SEX-LETHAL PROTEIN (C- TERMINUS, OR SECOND RNA-BINDING DOMAIN ISXL 3 (RBD-2), RESIDUES 199 - 294 PLUS N- TERMINAL MET) ISXL 4 (NMR, 17 STRUCTURES) ISXL 5	
377	2sxl		107	194	5.1e-22			57.11	SEX-LETHAL PROTEIN; CHAIN: NULL;	RNA-BINDING DOMAIN RNA-BINDING DOMAIN, ALTERNATIVE SPLICING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
377	2sxl		110	192	5.1e-22	0.90	1.00		SEX-LETHAL PROTEIN; CHAIN: NULL;	RNA-BINDING DOMAIN RNA-BINDING DOMAIN, ALTERNATIVE SPLICING
377	2up1	A	104	193	6.8e-32	1.01	1.00		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1; CHAIN: A; 12- NUCLEOTIDE SINGLE- STRANDED TELOMETRIC DNA; CHAIN: B;	COMPLEX (RIBONUCLEOPROTEIN/DN A) HNRNP A1, UP1; COMPLEX (RIBONUCLEOPROTEIN/DN A), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1
377	3sxl	A	26	179	1.7e-21			50.80	SEX-LETHAL; CHAIN: A, B, C;	RNA BINDING DOMAIN RNA BINDING DOMAIN, RBD, RNA RECOGNITION MOTIF, RRM, 2 SPLICING INHIBITOR, TRANSLATIONAL INHIBITOR, SEX 3 DETERMINATION, X CHROMOSOME DOSAGE COMPENSATION
380	1ehd	A	43	120	1.7e-05	-0.05	0.21		AGGLUTININ ISOLECTIN VI; CHAIN: A	PLANT PROTEIN TWO HOMOLOGOUS HEVEIN- LIKE DOMAINS
380	1eis	A	43	120	1.7e-05	-0.30	0.48		AGGLUTININ ISOLECTIN VI/AGGLUTININ ISOLECTIN V; CHAIN: A;	SUGAR BINDING PROTEIN UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPERANTIGEN
380	1en2	A	43	120	1e-05	0.38	0.45		AGGLUTININ ISOLECTIN I/AGGLUTININ ISOLECTIN V/ CHAIN: A;	SUGAR BINDING PROTEIN UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPERANTIGEN, SACCHARIDE BINDING
380	1ext	A	45	192	1.6e-07	-0.33	0.06		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
380	1hia	I	39	77	0.00033	-0.56	0.24		KALLIKREIN; CHAIN: A, B,	COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Pst Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									X, Y; HIRUSTASIN; CHAIN: I, J;	(PROTEASE/INHIBITOR) COMPLEX (PROTEASE/INHIBITOR), TISSUE KALLIKREIN, SERINE 2 PROTEASE, TRYPSIN, PSA, KININ, SERPIN
380	1klo		3	140	8.5e-09	0.40	-0.02		LAMININ; CHAIN: NULL;	GLYCOPROTEIN
380	1klo		43	204	3.3e-11	0.10	0.11		LAMININ; CHAIN: NULL;	GLYCOPROTEIN
380	1klo		43	207	3.3e-11			73.03	LAMININ; CHAIN: NULL;	GLYCOPROTEIN
380	1skz		39	138	1.7e-06	-0.21	0.18		ANTISTASIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR; ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2 SERINE PROTEASE INHIBITOR, THROMBOSIS
380	9wga	A	41	199	8.5e-09			60.96	LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
380	9wga	A	7	179	8.5e-09	-0.00	-0.19		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
385	1am4	D	31	225	9.9e-49			66.17	P50-RHOGAP; CHAIN: A, B, C; CDC42HS; CHAIN: D, E, F;	COMPLEX (GTPASE-ACTIVATING/GTP-BINDING) COMPLEX (GTPASE-ACTIVATING/GTP-BINDING), GTPASE ACTIVATION
385	1byu	A	29	227	1.3e-58			97.94	GTP-BINDING PROTEIN TC4; RAN; CHAIN: A, B;	TRANSPORT PROTEIN TC4; GTPASE, NUCLEAR TRANSPORT, TRANSPORT PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
385	1byu	A	32	211	1.3e-58	0.58	1.00		GTP-BINDING PROTEIN RAN; CHAIN: A, B;	TRANSPORT PROTEIN TC4; GTPASE, NUCLEAR TRANSPORT, TRANSPORT PROTEIN
385	1byu	B	26	227	9.9e-59			93.32	GTP-BINDING PROTEIN RAN; CHAIN: A, B;	TRANSPORT PROTEIN TC4; GTPASE, NUCLEAR TRANSPORT, TRANSPORT PROTEIN
385	1byu	B	27	211	9.9e-59	0.25	0.98		GTP-BINDING PROTEIN RAN; CHAIN: A, B;	TRANSPORT PROTEIN TC4; GTPASE, NUCLEAR TRANSPORT, TRANSPORT PROTEIN
385	1cly	A	31	201	3.4e-65			101.45	RAS-RELATED PROTEIN RAP-1A; CHAIN: A; PROTO-ONCOGENE SERINE/THREONINE PROTEIN KINASE CHAIN: B;	SIGNALING PROTEIN GTP-BINDING PROTEINS, PROTEIN-PROTEIN COMPLEX, EFFECTORS
385	1cly	A	32	201	3.4e-65	0.63	1.00		RAS-RELATED PROTEIN RAP-1A; CHAIN: A; PROTO-ONCOGENE SERINE/THREONINE PROTEIN KINASE CHAIN: B;	SIGNALING PROTEIN GTP-BINDING PROTEINS, PROTEIN-PROTEIN COMPLEX, EFFECTORS
385	1ctq	A	33	201	3.4e-66	0.53	1.00		TRANSFORMING PROTEIN P21/H-RAS-1; CHAIN: A;	SIGNALING PROTEIN G PROTEIN, GTP HYDROLYSIS, KINETIC CRYSTALLOGRAPHY, 2 SIGNALING PROTEIN
385	1ctq	A	33	202	3.4e-66			113.07	TRANSFORMING PROTEIN P21/H-RAS-1; CHAIN: A;	SIGNALING PROTEIN G PROTEIN, GTP HYDROLYSIS, KINETIC CRYSTALLOGRAPHY, 2 SIGNALING PROTEIN
385	1ctz	A	28	201	6.8e-51			87.72	HIS-TAGGED TRANSFORMING PROTEIN RHOA(0-181); CHAIN: A; PKN; CHAIN: B;	SIGNALING PROTEIN PROTEIN-PROTEIN COMPLEX, ANTIPARALLEL COILED-COIL
385	1d5c	A	34	199	5.1e-59	0.86	1.00		RAB6 GTPASE; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
385	1ds6	A	32	204	3.3e-53	0.49	1.00		RAS-RELATED C3 BOTULINUM TOXIN SUBSTRATE 2; CHAIN: A; RHO GDP-DISSOCIATION INHIBITOR 2; CHAIN: B;	S G-PROTEIN, GTPASE, RAB6, VESICULAR TRAFFICKING
385	1ek0	A	34	199	1e-56	0.90	1.00		GTP-BINDING PROTEIN YPT51; CHAIN: A;	SIGNALING PROTEIN P21-RAC2; RHO GDI 2, RHO-GDI BETA, LY-GDI; BETA SANDWICH, PROTEIN-PROTEIN COMPLEX, G-DOMAIN, 2 IMMUNOGLOBULIN FOLD, WALKER FOLD, GTP-BINDING PROTEIN
385	1hur	A	21	203	1.7e-11			66.72	HUMAN ADP-RIBOSYLATION FACTOR 1; 1HUR 5 CHAIN: A, B; 1HUR 7	ENDOCYTOSIS/EXOCYTOSIS G PROTEIN, VESICULAR TRAFFIC, GTP HYDROLYSIS, YPT/RAB 2 PROTEIN, ENDOCYTOSIS, HYDROLASE
385	1ibr	A	33	207	1.7e-51			104.05	RAN; CHAIN: A, C; IMPORTIN BETA SUBUNIT; CHAIN: B, D;	PROTEIN TRANSPORT GDP-BINDING, MEMBRANE TRAFFICKIN, NON-MYRISTOYLATED 1HUR 16
385	1kao		31	202	1.7e-60			104.05	RAP2A; CHAIN: NULL;	SMALL GTPASE KARYOPHERIN BETA, P95 SMALL GTPASE, NUCLEAR TRANSPORT RECEPTOR
385	1kao		32	199	1.7e-60	0.52	1.00		RAP2A; CHAIN: NULL;	GTP-BINDING PROTEIN GTP-BINDING PROTEIN, SMALL G PROTEIN, RAP2, GDP, RAS
385	1mh1		29	211	1.2e-52			87.53	RAC1; CHAIN: NULL;	GTP-BINDING PROTEIN GTP-BINDING PROTEIN, SMALL G PROTEIN, RAP2, GDP, RAS
385	1plj		33	201	6.8e-51			57.70	ONCOGENE PROTEIN C-H-	GTP-BINDING GTP-BINDING, GTPASE, SMALL G-PROTEIN, RHO FAMILY, RAS SUPER 2 FAMILY

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
385	1rrp	C	31	221	8.5e-52				RAS P21 PROTEIN MUTANT WITH GLY 12 REPLACED BY PRO 1PLJ 3 (G12P) COMPLEXED WITH P3-1-(2-NITROPHENYL)ETHYL-1PLJ 4 GUANOSINE-5'-(B,G-IMIDO)-TRIPHOSPHATE 1PLJ 5	
385	1tx4	B	31	199	5.1e-48			104.86	RAN; CHAIN: A, C; NUCLEAR PORE COMPLEX PROTEIN NUP358; CHAIN: B, D;	COMPLEX (SMALL GTPASE/NUCLEAR PROTEIN) COMPLEX (SMALL GTPASE/NUCLEAR PROTEIN), SMALL GTPASE, 2 NUCLEAR TRANSPORT
385	1zbd	A	29	203	1.7e-64			73.28	P50-RHO GAP; CHAIN: A; TRANSFORMING PROTEIN RHOA; CHAIN: B;	COMPLEX (GTPASE ACTIVATION/PROTO-ONCOGENE) GTPASE-ACTIVATING PROTEIN RHO GAP; COMPLEX (GTPASE ACTIVATION/PROTO-ONCOGENE), GTPASE, 2 TRANSITION STATE, GAP
385	1zbd	A	30	206	1.7e-64	0.76	1.00	115.76	RAB-3A; CHAIN: A; RABPHILIN-3A; CHAIN: B;	COMPLEX (GTP-BINDING/EFFECTOR) RAS-RELATED PROTEIN RAB3A, COMPLEX (GTP-BINDING/EFFECTOR), G PROTEIN, EFFECTOR, RABCDR, 2 SYNAPTIC EXOCYTOSIS, RAB PROTEIN, RAB3A, RABPHILIN
385	1zbd	A	30	206	1.7e-64	0.76	1.00	115.76	RAB-3A; CHAIN: A; RABPHILIN-3A; CHAIN: B;	COMPLEX (GTP-BINDING/EFFECTOR) RAS-RELATED PROTEIN RAB3A, COMPLEX (GTP-BINDING/EFFECTOR), G

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
385	2ngr	A	31	229	9.9e-53			79.16	GTP BINDING PROTEIN (G25K); CHAIN: A; GTPASE ACTIVATING PROTEIN (RHG); CHAIN: B;	HYDROLASE CDC42/CDC42GAP; CDC42/CDC42GAP; TRANSITION STATE, G- PROTEIN, GAP, CDC42, ALF3, HYDROLASE HYDROLASE G PROTEIN, VESICULAR TRAFFICKING, GTP HYDROLYSIS, RAB 2 PROTEIN, NEUROTTRANSMITTER RELEASE, HYDROLASE HYDROLASE G PROTEIN, VESICULAR TRAFFICKING, GTP HYDROLYSIS, RAB 2 PROTEIN, NEUROTTRANSMITTER RELEASE, HYDROLASE
385	3rab	A	29	202	3.4e-65	0.89	1.00		RAB3A; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS S SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
385	3rab	A	30	202	3.4e-65			127.86	RAB3A; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS S SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
386	1ez3	A	195	292	3.3e-08	0.00	-0.18		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS S SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
386	1ez3	A	94	227	2.3e-10	0.03	-0.17		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS S SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
386	1quu	A	21	292	3.3e-18	0.04	-0.07		HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
388	1a4y	A	13	201	8.5e-05	-0.16	0.37		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
388	1a4y	A	46	276	8.5e-14	0.19	0.99		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
388	1a4y	A	93	276	2.6e-17	0.44	1.00		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
388	1a9n	A	94	227	0.0037	0.21	0.98		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
388	1a9n	C	27	175	0.0066	0.29	0.05		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
388	1a9n	C	94	227	0.0037	0.09	0.96		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
388	1d0b	A	150	275	8.5e-09	-0.38	0.03		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
388	1d0b	A	88	254	6.8e-14	-0.22	0.07		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
388	1ds9	A	135	276	1e-07	0.07	0.13		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
388	1ds9	A	91	227	1e-07	0.17	0.17		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
388	1fo1	A	146	253	1.7e-07	0.05	0.41		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP, RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
388	1fo1	A	81	227	1.7e-08	0.12	0.13		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP, RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
388	1fo1	B	81	227	1.7e-08	-0.02	0.09		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
388	1fqv	A	25	276	3.4e-33	0.25	1.00		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	(RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
388	1fqv	A	8	198	1.7e-11	0.28	0.24		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN
388	1fs1	A	25	65	1e-07	-0.61	0.70		CYCLIN A/CDK2-ASSOCIATED P19; CHAIN: A, C; CYCLIN A/CDK2-ASSOCIATED P45; CHAIN: B, D;	LIGASE SKP2 F-BOX; SKP1; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN
388	1fs2	A	25	276	5.1e-35	0.68	1.00		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN
388	1fs2	A	8	198	1.7e-11	0.23	0.39		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, UBIQUITIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
388	2bnh		13	227	1.2e-12	0.15	0.62		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
388	2bnh		2	272	1.4e-11	0.20	0.71		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
388	2bnh		46	276	1e-15	0.35	0.95		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
389	1ysa	C	73	127	6.6e-16	0.08	-0.19		LEUCINE ZIPPER GCN4 (BASIC REGION, LEUCINE ZIPPER) COMPLEX WITH AP-1 DNA 1YSA 3	
392	1alh	A	155	225	5.1e-27	0.01	0.92		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
392	1alh	A	173	253	1.7e-29	0.13	1.00		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
392	1alh	A	173	255	1.7e-29			72.18	C; QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLICATION OF OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
392	1alh	A	201	264	1.7e-24	-0.07	0.51		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLICATION OF OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
392	1mey	C	154	225	3.4e-41	-0.20	0.65		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
392	1mey	C	172	253	1.7e-50	0.13	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
392	1mey	C	172	254	1.7e-50			83.29	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
392	1mey	C	200	264	1.2e-40	0.11	0.98		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
392	1mey	G	228	253	3.3e-09	0.46	0.86		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	STRUCTURE, COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
392	1ef3	A	162	225	6.8e-16	-0.01	0.40		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
392	1ef3	A	172	263	1.2e-19			56.76	TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
392	1ef3	A	173	253	1.2e-19	-0.03	0.87		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
392	1tf3	A	201	266	1.7e-17	-0.12	0.43		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
392	1tf6	A	162	262	1.7e-27	-0.39	0.01		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
392	1tf6	A	72	257	1.7e-27			53.26	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
392	1ubd	C	140	254	6.8e-31			66.82	YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
392	1ubd	C	156	253	6.8e-31	-0.06	0.98		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
392	1ubd	C	180	261	1.7e-28	-0.04	0.60		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
392	2gli	A	152	252	5.1e-29	-0.16	0.94		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING, PROTEIN/DNA)
392	2gli	A	99	254	5.1e-29			58.63	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING, PROTEIN/DNA)
393	1alh	A	669	747	1e-17	0.37	-0.20		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
393	1alh	A	702	781	5.1e-22	0.42	-0.11		QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
393	1alh	A	723	809	6.8e-27	0.35	0.21		QGSZ ZINC FINGER PEPTIDE; CHAIN: A;	COMPLEX (ZINC FINGER/DNA) COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
393	1alh	A	757	833	8.5e-28	0.33	0.78		DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	(ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
393	1bbo		759	813	2e-14	-0.28	0.51		QQR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
393	1chc		11	35	0.00013	-0.77	0.25		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11. IBBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) IBBO 4	
393	1chc		11	56	0.0097	-0.44	0.11		VIRUS EQUINE HERPES VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	
393	1fbv	A	11	52	0.0097	-0.29	0.18		VIRUS EQUINE HERPES VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	LIGASE CBL, UBCH7, ZAP-70, E2, UBIQUITIN, E3, PHOSPHORYLATION, 2 TYROSINE KINASE, UBIQUITINATION, PROTEIN DEGRADATION,
393	1g25	A	11	63	6.6e-05	-0.13	0.18		SIGNAL TRANSDUCTION PROTEIN CBL; CHAIN: A; ZAP-70 PEPTIDE; CHAIN: B; UBIQUITIN-CONJUGATING ENZYME E12-18 KDA UBCH7; CHAIN: C;	METAL BINDING PROTEIN RING FINGER PROTEIN MAT1; RING FINGER (C3HC4)
393	1mey	C	695	781	3.4e-40	0.29	-0.13		CDK-ACTIVATING KINASE ASSEMBLY FACTOR MAT1; CHAIN: A;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER PROTEIN-DNA

Table 5

SEQ ID NO;	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
393	1mey	C	722	809	5.1e-49	0.37	0.64		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
393	1mey	C	756	835	5.1e-50	0.58	0.99		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
393	1mey	G	782	809	1.5e-13	0.08	0.95		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
393	1rmd		3	35	0.00051	-0.66	0.33		RAG1; CHAIN: NULL;	DNA-BINDING PROTEIN V(D)J RECOMBINATION ACTIVATING PROTEIN 1; RAG1, V(D)J RECOMBINATION, ANTIBODY, MAD, RING FINGER, 2 ZINC BINUCLEAR CLUSTER, ZINC FINGER, DNA-BINDING PROTEIN
393	1tf3	A	757	833	3.4e-19	0.29	0.62		TRANSCRIPTION FACTOR IIIa; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA, 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, R

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
393	1tf6	A	647	762	5.1e-13	0.14	-0.20		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE II, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
393	1tf6	A	657	790	5.1e-21	0.18	-0.19		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE II, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
393	1tf6	A	668	818	3.4e-29	0.00	-0.17		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE II, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
393	1tf6	A	723	835	3.4e-26	0.22	-0.15		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE II, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
393	1ubd	C	647	748	1.7e-19	0.07	-0.20		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING- F1

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									INITIATOR ELEMENT DNA; CHAIN: A, B;	YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
393	1ubd	C	675	781	1e-23	0.28	-0.17		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
393	1ubd	C	698	809	1e-29	0.33	-0.01		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
393	1ubd	C	730	833	3.4e-32	0.37	0.22		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
393	2adr		757	811	5.1e-16	0.30	0.76		ADR1; CHAIN: NULL;	TRANSCRIPTION REGULATION, ADR1, ZINC 2 FINGER, NMR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
393	2adr		787	833	8.5e-14	-0.28	0.11		ADR1; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADR1, ZINC FINGER, NMR
393	2drp	A	757	811	9.9e-16	0.33	0.94		COMPLEX(TRANSCRIPTION REGULATION/DNA) TRAMTRACK PROTEIN (TWO ZINC-FINGER PEPTIDE) COMPLEXED WITH 2DRP 3 DNA 2DRP 4	
393	2gli	A	650	780	1.2e-22	0.25	-0.20		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
393	2gli	A	672	808	1.7e-27	0.12	-0.12		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
393	2gli	A	702	833	1.2e-30	0.29	0.30		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
394	1buo	A	18	141	3.4e-22	0.22	1.00		PROMYELOCYTIC LEUKEMIA ZINC FINGER PROTEIN PLZF; CHAIN: A;	GENE REGULATION POZ DOMAIN; PROTEIN-PROTEIN INTERACTION DOMAIN, TRANSCRIPTIONAL 2 REPRESSOR, ZINC-FINGER PROTEIN, X-RAY CRYSTALLOGRAPHY, 3 PROTEIN STRUCTURE, PROMYELOCYTIC LEUKEMIA, GENE REGULATION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
394	1buo	A	18	144	3.4e-22			73.07	PROMYELOCYTIC LEUKEMIA ZINC FINGER PROTEIN PLZF; CHAIN: A;	GENE REGULATION POZ DOMAIN; PROTEIN-PROTEIN INTERACTION DOMAIN, TRANSCRIPTIONAL 2 REPRESSOR, ZINC-FINGER PROTEIN, X-RAY CRYSTALLOGRAPHY, 3 PROTEIN STRUCTURE, PROMYELOCYTIC LEUKEMIA, GENE REGULATION
394	1gof		306	587	1e-17	0.13	0.76		OXIDOREDUCTASE(OXYGEN(A)) GALACTOSE OXIDASE (E.C.1.1.3.9) (PH 4.5) 1GOF 3	
395	1byn	A	141	274	1.3e-40	0.88	1.00		SYNAPTOTAGMIN I; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN, C2-DOMAIN, EXOCYTOSIS, NEUROTRANSMITTER 2 RELEASE, ENDOCYTOSIS/EXOCYTOSIS
395	1byn	A	146	269	3.4e-34	0.74	1.00		SYNAPTOTAGMIN I; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN, C2-DOMAIN, EXOCYTOSIS, NEUROTRANSMITTER 2 RELEASE, ENDOCYTOSIS/EXOCYTOSIS
395	1dqv	A	141	418	1.7e-87	0.60	1.00		SYNAPTOTAGMIN III; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS S BETA SANDWICH, CALCIUM ION, C2 DOMAIN
395	1dqv	A	146	417	1.5e-57	0.48	1.00		SYNAPTOTAGMIN III; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS S BETA SANDWICH, CALCIUM ION, C2 DOMAIN
395	1dsy	A	296	427	1e-20	0.19	0.36		PROTEIN KINASE C, ALPHA TYPE; CHAIN: A;	TRANSFERASE CALCIUM++ PHOSPHOLIPID BINDING

Table 5

SEQ ID NO.	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
395	lrsy		132	271	9.9e-46	0.70	1.00		CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN I (FIRST C2 DOMAIN) (CALB) IRSY 3	PROTEIN, CALCIUM-BINDING 2 PROTEIN, PHOSPHATIDYL SERINE, PROTEIN KINASE C
395	lrsy		132	272	9.9e-46			169.67	CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN I (FIRST C2 DOMAIN) (CALB) IRSY 3	
395	lrsy		146	269	3.4e-34	0.79	1.00		CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN I (FIRST C2 DOMAIN) (CALB) IRSY 3	
396	lase		10	124	6.6e-23	0.20	1.00		TUMOR SUPPRESSOR P16INK4A; CHAIN: NULL;	ANTI-ONCOGENE CELL CYCLE, ANTI-ONCOGENE, REPEAT, ANK REPEAT
396	lawc	B	180	323	5.1e-34	-0.10	0.03		GA BINDING PROTEIN ALPHA; CHAIN: A; GA BINDING PROTEIN BETA 1; CHAIN: B; DNA; CHAIN: D, E;	COMPLEX (TRANSCRIPTION REGULATION/DNA) GABP ALPHA; GABP BETA 1; COMPLEX (TRANSCRIPTION REGULATION/DNA), DNA-BINDING, 2 NUCLEAR PROTEIN, ETS DOMAIN, ANKYRIN REPEATS, TRANSCRIPTION 3 FACTOR
396	lawc	B	52	224	1.2e-33	-0.13	1.00		GA BINDING PROTEIN ALPHA; CHAIN: A; GA BINDING PROTEIN BETA 1; CHAIN: B; DNA; CHAIN: D, E;	COMPLEX (TRANSCRIPTION REGULATION/DNA) GABP ALPHA; GABP BETA 1; COMPLEX (TRANSCRIPTION REGULATION/DNA), DNA-BINDING, 2 NUCLEAR PROTEIN, ETS DOMAIN,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
396	1awc	B	9	195	6.8e-35	0.16	1.00		GA BINDING PROTEIN ALPHA; CHAIN: A; GA BINDING PROTEIN BETA 1; CHAIN: B; DNA; CHAIN: D, E;	ANKYRIN REPEATS, TRANSCRIPTION 3 FACTOR COMPLEX (TRANSCRIPTION REGULATION/DNA) GABP ALPHA; GABP BETA 1; COMPLEX (TRANSCRIPTION REGULATION/DNA), DNA-BINDING, 2 NUCLEAR PROTEIN, ETS DOMAIN, ANKYRIN REPEATS, TRANSCRIPTION 3 FACTOR
396	1bd8		10	198	1.3e-30	0.14	0.88		P19INK4D CDK4/6 INHIBITOR; CHAIN: NULL;	TUMOR SUPPRESSOR TUMOR SUPPRESSOR, CDK4/6 INHIBITOR, ANKYRIN MOTIF
396	1bd8		12	195	1e-27	0.09	1.00		P19INK4D CDK4/6 INHIBITOR; CHAIN: NULL;	TUMOR SUPPRESSOR TUMOR SUPPRESSOR, CDK4/6 INHIBITOR, ANKYRIN-MOTIF
396	1bd8		48	226	9.9e-31	0.21	0.81		P19INK4D CDK4/6 INHIBITOR; CHAIN: NULL;	TUMOR SUPPRESSOR TUMOR SUPPRESSOR, CDK4/6 INHIBITOR, ANKYRIN MOTIF
396	1bi7	B	10	102	1.6e-20	0.25	1.00		CYCLIN-DEPENDENT KINASE 6; CHAIN: A; MULTIPLE TUMOR SUPPRESSOR; CHAIN: B;	COMPLEX (KINASE/ANTI-ONCOGENE) CDK6; P16INK4A, MTS1; CYCLIN DEPENDENT KINASE, CYCLIN DEPENDENT KINASE INHIBITORY 2 PROTEIN, CDK, INK4, CELL CYCLE, MULTIPLE TUMOR SUPPRESSOR, 3 MTS1, COMPLEX (KINASE/ANTI-ONCOGENE) HEADER
396	1blx	B	12	195	1.2e-26	0.28	1.00		CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
396	1blx	B	50	226	3e-28	0.06	0.21		CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
396	1blx	B	5	202	9.9e-34	0.29	1.00			COMPLEX (INHIBITOR PROTEIN/KINASE)
396	1bu9	A	9	200	1.7e-32	0.08	0.39		CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A;	COMPLEX (INHIBITOR PROTEIN/KINASE)
396	1hbb	A	183	322	1.5e-28	0.06	-0.17		CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A, B;	HORMONE/GROWTH FACTOR P18-INK4C; CELL CYCLE INHIBITOR, P18INK4C, TUMOR, SUPPRESSOR, CYCLIN-2 DEPENDENT KINASE, HORMONE/GROWTH FACTOR
396	1hbb	A	9	199	1e-31	0.27	1.00		CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A, B;	CELL CYCLE INHIBITOR P18-INK4C(INK6); CELL CYCLE INHIBITOR, P18-INK4C(INK6), ANKYRIN REPEAT, 2 CDK 4/6 INHIBITOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
396	1lkn	D	4	235	6.8e-38	-0.06	0.94		NF-KAPPA-B P65 SUBUNIT; CHAIN: A; NF-KAPPA-B P50D SUBUNIT; CHAIN: C; I-KAPPA-B-ALPHA; CHAIN: D;	TRANSCRIPTION FACTOR P65; P50D; TRANSCRIPTION FACTOR, IKB/NFKB COMPLEX
396	1myo		10	135	1.4e-23	-0.02	1.00		MYOTROPHIN; CHAIN: NULL	ANK-REPEAT MYOTROPHIN, ACETYLATION, NMR, ANK-REPEAT
396	1myo		53	164	1.2e-22	-0.16	0.53		MYOTROPHIN; CHAIN: NULL	ANK-REPEAT MYOTROPHIN, ACETYLATION, NMR, ANK-REPEAT
396	1mf	B	1	235	3.4e-39	-0.05	0.77		NF-KAPPA-B P65; CHAIN: A, C; NF-KAPPA-B P50; CHAIN: B, D; I-KAPPA-B-ALPHA; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REG/ANK REPEAT) COMPLEX (TRANSCRIPTION REGULATION/ANK REPEAT), ANKYRIN 2 REPEAT HELIX
396	1sw6	A	1	218	2e-32	-0.08	0.99		REGULATORY PROTEIN SWI6; CHAIN: A, B;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ANKYRIN REPEATS, CELL-CYCLE
396	1ycs	B	9	217	1.3e-25	-0.27	0.54		P53; CHAIN: A; 53BP2; CHAIN: B;	COMPLEX (ANTI-ONCOGENE/ANKYRIN REPEATS) P53BP2; ANKYRIN REPEATS, SH3, P53, TUMOR SUPPRESSOR, MULTIGENE 2 FAMILY, NUCLEAR PROTEIN, PHOSPHORYLATION, DISEASE MUTATION, 3 POLYMORPHISM, COMPLEX (ANTI-ONCOGENE/ANKYRIN REPEATS)
397	1awc	B	16	152	8.5e-39			66.19	GA BINDING PROTEIN	COMPLEX (TRANSCRIPTION)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									ALPHA; CHAIN: A; GA BINDING PROTEIN BETA 1; CHAIN: B; DNA; CHAIN: D, E;	REGULATION/DNA) GABPALPHA; GABPBETA1; COMPLEX (TRANSCRIPTION REGULATION/DNA), DNA-BINDING, 2 NUCLEAR PROTEIN, ETS DOMAIN, ANKYRIN REPEATS, TRANSCRIPTION 3 FACTOR
397	lawc	B	1	136	8.5e-39	0.64	1.00		GA BINDING PROTEIN ALPHA; CHAIN: A; GA BINDING PROTEIN BETA 1; CHAIN: B; DNA; CHAIN: D, E;	COMPLEX (TRANSCRIPTION REGULATION/DNA) GABPALPHA; GABPBETA1; COMPLEX (TRANSCRIPTION REGULATION/DNA), DNA-BINDING, 2 NUCLEAR PROTEIN, ETS DOMAIN, ANKYRIN REPEATS, TRANSCRIPTION 3 FACTOR
397	lawc	B	22	152	6.8e-37	0.68	1.00		GA BINDING PROTEIN ALPHA; CHAIN: A; GA BINDING PROTEIN BETA 1; CHAIN: B; DNA; CHAIN: D, E;	COMPLEX (TRANSCRIPTION REGULATION/DNA) GABPALPHA; GABPBETA1; COMPLEX (TRANSCRIPTION REGULATION/DNA), DNA-BINDING, 2 NUCLEAR PROTEIN, ETS DOMAIN, ANKYRIN REPEATS, TRANSCRIPTION 3 FACTOR
397	lbd8		1	139	1.5e-31	0.55	1.00		P19INK4D CDK4/6 INHIBITOR; CHAIN: NULL;	TUMOR SUPPRESSOR, CDK4/6 INHIBITOR, ANKYRIN MOTIF
397	lbu9	B	1	139	5.1e-30	0.38	1.00		CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
397	lbu9	A	1	149	1.2e-35			56.15	CYCLIN-DEPENDENT	HORMONE/GROWTH

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									KINASE 6 INHIBITOR; CHAIN: A;	FACTOR P18-INK4C; CELL CYCLE INHIBITOR, P18INK4C, TUMOR, SUPPRESSOR, CYCLIN-2 DEPENDENT KINASE, HORMONE/GROWTH FACTOR
397	1bu9	A	4	141	1.2e-35	0.40	1.00		CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A;	HORMONE/GROWTH FACTOR P18-INK4C; CELL CYCLE INHIBITOR, P18INK4C, TUMOR, SUPPRESSOR, CYCLIN-2 DEPENDENT KINASE, HORMONE/GROWTH FACTOR
397	1ihb	A	1	140	5.1e-35			56.08	CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A, B;	CELL CYCLE INHIBITOR P18-INK4C(INK6); CELL CYCLE INHIBITOR, P18-INK4C(INK6), ANKYRIN REPEAT, 2 CDK 4/6 INHIBITOR
397	1ihb	A	4	140	5.1e-35	0.51	1.00		CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A, B;	CELL CYCLE INHIBITOR P18-INK4C(INK6); CELL CYCLE INHIBITOR, P18-INK4C(INK6), ANKYRIN REPEAT, 2 CDK 4/6 INHIBITOR
397	1ikn	D	2	120	1.7e-30	0.37	1.00		NF-KAPPA-B P65 SUBUNIT; CHAIN: A; NF-KAPPA-B P50D SUBUNIT; CHAIN: C; I-KAPPA-B-ALPHA; CHAIN: D;	TRANSCRIPTION FACTOR P65; P50D; TRANSCRIPTION FACTOR, IKB/NFKB COMPLEX
397	1ikn	D	9	152	1e-38	0.21	0.81		NF-KAPPA-B P65 SUBUNIT; CHAIN: A; NF-KAPPA-B P50D SUBUNIT; CHAIN: C; I-KAPPA-B-ALPHA; CHAIN: D;	TRANSCRIPTION FACTOR P65; P50D; TRANSCRIPTION FACTOR, IKB/NFKB COMPLEX
397	1myo		17	134	1.7e-27			52.07	MYOTROPHIN; CHAIN: NULL	ANK-REPEAT MYOTROPHIN,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
397	1nfi	E	2	120	1.7e-30	0.59	1.00		NF-KAPPA-B P65; CHAIN: A, C; NF-KAPPA-B P50; CHAIN: B, D; I-KAPPA-B-ALPHA; CHAIN: E, F; REPEAT, ANKYRIN 2 REPEAT HELIX	ACETYLATION, NMR, ANK-REPEAT
397	1nfi	E	9	152	1e-38	0.51	1.00		NF-KAPPA-B P65; CHAIN: A, C; NF-KAPPA-B P50; CHAIN: B, D; I-KAPPA-B-ALPHA; CHAIN: E, F; REPEAT, ANKYRIN 2 REPEAT HELIX	COMPLEX (TRANSCRIPTION REG/ANK REPEAT)
400	1cdm	A	91	239	6.8e-33	-0.27	0.37		CALCIUM-BINDING PROTEIN CALMODULIN COMPLEXED WITH CALMODULIN-BINDING DOMAIN OF ICDM 3 CALMODULIN-DEPENDENT PROTEIN KINASE II ICDM 4	COMPLEX (TRANSCRIPTION REG/ANK REPEAT)
400	1cll		91	239	1.2e-36	-0.42	0.35		CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) ICDM 3	COMPLEX (TRANSCRIPTION REG/ANK REPEAT)
400	1exr	A	89	239	6.8e-35	-0.42	0.15		CALMODULIN; CHAIN: A;	METAL TRANSPORT CALMODULIN, HIGH RESOLUTION, DISORDER
400	1faq		261	310	1.2e-06	0.33	0.58		RAF-1; CHAIN: NULL;	SERINE/THREONINE PROTEIN KINASE TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, 2 PROTO ONCOGENE, ZINC, ATP-BINDING, PHORBOL-ESTER BINDING
400	1faq		261	312	2.6e-15	0.19	0.47		RAF-1; CHAIN: NULL;	SERINE/THREONINE PROTEIN KINASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
400	1faq		334	374	2e-07	-0.42	0.01		RAF-1; CHAIN: NULL;	SERINE/THREONINE-PROTEIN KINASE, 2 PROTO-ONCOGENE, ZINC, ATP-BINDING, PHORBOL-ESTER BINDING
400	1faq		336	374	6.8e-06	-0.37	0.19		RAF-1; CHAIN: NULL;	SERINE/THREONINE-PROTEIN KINASE, 2 PROTO-ONCOGENE, ZINC, ATP-BINDING, PHORBOL-ESTER BINDING
400	1ptq		261	310	6.6e-17	-0.16	0.19		PROTEIN KINASE C DELTA TYPE; IPTQ 4	PHOSPHOTRANSFERASE
400	1ptq		261	310	6.8e-10	-0.16	0.19		PROTEIN KINASE C DELTA TYPE; IPTQ 4	PHOSPHOTRANSFERASE
400	1ptq		325	374	3.4e-17	0.17	0.31		PROTEIN KINASE C DELTA TYPE; IPTQ 4	PHOSPHOTRANSFERASE
400	1tbn		261	310	1e-09	-0.36	0.25		PROTEIN KINASE C, GAMMA TYPE; CHAIN: NULL;	CALCIUM-BINDING PROTEIN RAT BRAIN PKC-G; CALCIUM-BINDING PROTEIN, PROTEIN KINASE C, PKC, TRANSFERASE
400	1tbn		261	316	1.7e-18	-0.22	0.25		PROTEIN KINASE C, GAMMA TYPE; CHAIN: NULL;	CALCIUM-BINDING PROTEIN RAT BRAIN PKC-G; CALCIUM-BINDING PROTEIN, PROTEIN KINASE C, PKC, TRANSFERASE
400	1tnx		91	239	1.7e-25	-0.56	0.11		TROPONIN C; ITNX 4	CALCIUM-BINDING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
400	1top		91	239	3.4e-27	0.07	0.06		CHAIN: NULL; ITNX 5 CONTRACTILE SYSTEM PROTEIN TROPONIN C ITOP 3	PROTEIN EF-HAND ITNX 14
400	1vrk	A	88	240	6.8e-35	-0.41	0.22		CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP- HELIX, SIGNALLING, 2 COMPLEX(CALCIUM- BINDING PROTEIN/PEPTIDE)
403	1b8q	A	222	290	3.3e-15	0.64	1.00		NEURONAL NITRIC OXIDE SYNTHASE; CHAIN: A; HEPTAPEPTIDE; CHAIN: B;	OXIDOREDUCTASE PDZ DOMAIN, NNOS, NITRIC OXIDE SYNTHASE
403	1b8q	A	818	893	9.9e-16	0.10	0.66		NEURONAL NITRIC OXIDE SYNTHASE; CHAIN: A; HEPTAPEPTIDE; CHAIN: B;	OXIDOREDUCTASE PDZ DOMAIN, NNOS, NITRIC OXIDE SYNTHASE
403	1b8q	A	84	157	9.9e-17	0.29	0.99		NEURONAL NITRIC OXIDE SYNTHASE; CHAIN: A; HEPTAPEPTIDE; CHAIN: B;	OXIDOREDUCTASE PDZ DOMAIN, NNOS, NITRIC OXIDE SYNTHASE
403	1be9	A	232	288	5.1e-07	0.26	0.99		PSD-95; CHAIN: A; CRPT; CHAIN: B;	PEPTIDE RECOGNITION, PEPTIDE RECOGNITION, PROTEIN LOCALIZATION
403	1be9	A	75	166	1.7e-16	0.61	1.00		PSD-95; CHAIN: A; CRPT; CHAIN: B;	PEPTIDE RECOGNITION, PEPTIDE RECOGNITION, PEPTIDE RECOGNITION, PROTEIN LOCALIZATION
403	1be9	A	812	900	1.4e-17	0.14	0.39		PSD-95; CHAIN: A; CRPT; CHAIN: B;	PEPTIDE RECOGNITION, PEPTIDE RECOGNITION, PEPTIDE RECOGNITION, PROTEIN LOCALIZATION
403	1kwa	A	222	293	3.3e-15	0.51	0.92		HCASK/LIN-2 PROTEIN; CHAIN: A, B;	KINASE HCASK, GLGF REPEAT, DHR; PDZ DOMAIN, NEUREXIN, SYNDECAN, RECEPTOR CLUSTERING, KINASE
403	1kwa	A	820	893	3.3e-14	-0.16	0.39		HCASK/LIN-2 PROTEIN; CHAIN: A, B;	KINASE HCASK, GLGF REPEAT, DHR; PDZ DOMAIN, NEUREXIN,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
403	1kwa	A	84	168	1.3e-18	0.41	0.94		HCASK/LIN-2 PROTEIN; CHAIN: A, B;	SYNDECAN, RECEPTOR CLUSTERING, KINASE
403	1pdr		222	273	9.9e-12	0.16	1.00		HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	KINASE HCASK, GLGF REPEAT, DHR; PDZ DOMAIN, NEUREXIN, SYNDECAN, RECEPTOR CLUSTERING, KINASE
403	1pdr		818	904	1.7e-16	0.05	0.34		HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	SIGNAL TRANSDUCTION HDLG, DHR3 DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, REPEAT
403	1pdr		83	166	1.7e-15	0.66	0.98		HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	SIGNAL TRANSDUCTION HDLG, DHR3 DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, REPEAT
403	1qau	A	222	303	6.6e-15	0.38	0.87		NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: A;	OXIDOREDUCTASE BETA-FINGER
403	1qau	A	84	170	2e-18	0.67	0.93		NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: A;	OXIDOREDUCTASE BETA-FINGER
403	1qav	A	222	289	6.6e-17	0.91	1.00		ALPHA-1 SYNTROPHIN (RESIDUES 77-171); CHAIN: A; NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: B;	MEMBRANE PROTEIN/OXIDOREDUCTASE BETA-FINGER, HETERODIMER
403	1qav	A	815	905	1.7e-16	-0.18	0.35		ALPHA-1 SYNTROPHIN (RESIDUES 77-171); CHAIN: A; NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: B;	MEMBRANE PROTEIN/OXIDOREDUCTASE BETA-FINGER, HETERODIMER

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
403	1qav	A	816	893	6.6e-15	0.53	1.00		ALPHA-1 SYNTROPHIN (RESIDUES 77-171); CHAIN: A; NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: B;	MEMBRANE PROTEIN/OXIDOREDUCTAS E BETA-FINGER, HETERODIMER
403	1qav	A	81	165	9.9e-20	0.64	1.00		ALPHA-1 SYNTROPHIN (RESIDUES 77-171); CHAIN: A; NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: B;	MEMBRANE PROTEIN/OXIDOREDUCTAS E BETA-FINGER, HETERODIMER
403	1qav	A	81	168	5.1e-18	0.77	0.99		ALPHA-1 SYNTROPHIN (RESIDUES 77-171); CHAIN: A; NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: B;	MEMBRANE PROTEIN/OXIDOREDUCTAS E BETA-FINGER, HETERODIMER
403	1qlc	A	222	289	2.3e-14	0.77	1.00		POSTSYNAPTIC DENSITY PROTEIN 95; CHAIN: A;	PEPTIDE RECOGNITION PSD-95; PDZ DOMAIN, NEURONAL NITRIC OXIDE SYNTHASE, NMDA RECEPTOR 2 BINDING
403	1qlc	A	816	893	2.3e-16	0.22	1.00		POSTSYNAPTIC DENSITY PROTEIN 95; CHAIN: A;	PEPTIDE RECOGNITION PSD-95; PDZ DOMAIN, NEURONAL NITRIC OXIDE SYNTHASE, NMDA RECEPTOR 2 BINDING
403	1qlc	A	819	901	1.2e-14	0.07	0.99		POSTSYNAPTIC DENSITY PROTEIN 95; CHAIN: A;	PEPTIDE RECOGNITION PSD-95; PDZ DOMAIN, NEURONAL NITRIC OXIDE SYNTHASE, NMDA RECEPTOR 2 BINDING
403	1qlc	A	81	165	1.3e-17	0.72	1.00		POSTSYNAPTIC DENSITY PROTEIN 95; CHAIN: A;	PEPTIDE RECOGNITION PSD-95; PDZ DOMAIN, NEURONAL NITRIC OXIDE SYNTHASE, NMDA RECEPTOR 2 BINDING
403	3pdz	A	222	289	1.6e-14	0.45	1.00		TYROSINE PHOSPHATASE	HYDROLASE PDZ DOMAIN, NMDA RECEPTOR 2 BINDING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
403	3pdz	A	233	296	8.5e-09	0.02	0.88		(PTP-BAS, TYPE 1); CHAIN: A; TYROSINE PHOSPHATASE (PTP-BAS, TYPE 1); CHAIN: A;	HUMAN PHOSPHATASE, HPTPIE, PTP-BAS, SPECIFICITY 2 OF BINDING HYDROLASE PDZ DOMAIN, HUMAN PHOSPHATASE, HPTPIE, PTP-BAS, SPECIFICITY 2 OF BINDING
403	3pdz	A	814	907	1.7e-16	0.26	0.41		TYROSINE PHOSPHATASE (PTP-BAS, TYPE 1); CHAIN: A;	HYDROLASE PDZ DOMAIN, HUMAN PHOSPHATASE, HPTPIE, PTP-BAS, SPECIFICITY 2 OF BINDING
403	3pdz	A	81	171	1.7e-15	0.37	0.76		TYROSINE PHOSPHATASE (PTP-BAS, TYPE 1); CHAIN: A;	HYDROLASE PDZ DOMAIN, HUMAN PHOSPHATASE, HPTPIE, PTP-BAS, SPECIFICITY 2 OF BINDING
403	3pdz	A	84	170	1.3e-16	0.27	1.00		TYROSINE PHOSPHATASE (PTP-BAS, TYPE 1); CHAIN: A;	HYDROLASE PDZ DOMAIN, HUMAN PHOSPHATASE, HPTPIE, PTP-BAS, SPECIFICITY 2 OF BINDING
404	1alh	A	147	222	1.7e-23	-0.55	0.68		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
404	1alh	A	170	250	3.4e-29	0.33	1.00		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
404	1alh	A	171	251	3.3e-33	-0.04	1.00		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
404	1alh	A	226	306	3.4e-31	0.49	1.00		QGR ZINC FINGER	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
404	1alh	A	226	307	9.9e-34	0.43	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
404	1alh	A	226	308	9.9e-34			80.11	QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
404	1alh	A	254	334	3.3e-32	0.12	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
404	1alh	A	282	362	9.9e-31	0.38	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
404	1alh	A	310	393	3.4e-19	-0.09	0.11		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
404	1mey	C	118	194	6.8e-36	-0.26	0.03		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, DNA-BINDING PROTEIN-INTERACTION, PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
404	1mey	C	146	222	1.7e-42	-0.06	0.86		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
404	1mey	C	169	250	1.7e-50	0.08	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
404	1mey	C	197	278	3.4e-51	0.35	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
404	1mey	C	225	306	6.8e-52	0.38	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
404	1mey	C	253	362	5.1e-47	0.32	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
404	1mey	C	281	363	5.1e-47			98.26	DNA; CHAIN: A, B, D, E;	COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
404	1mey	C	309	393	8.5e-38	-0.12	0.04		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
404	1tf3	A	225	309	1.7e-18			60.10	TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
404	1tf6	A	167	337	3.4e-38			114.59	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
404	1tf6	A	170	315	3.4e-38	-0.07	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
404	1uf6	A	254	403	5.1e-34	-0.04	0.62		TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
404	1ubd	C	148	250	1e-31	0.03	0.98		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
404	1ubd	C	174	278	1.3e-39	0.05	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
404	1ubd	C	177	278	1.2e-34	0.16	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
404	1ubd	C	223	335	3.3e-43	0.23	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
404	1ubd	C	225	335	3.3e-43			87.63	YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
404	1ubd	C	233	362	5.1e-32	0.04	0.89		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
404	1ubd	C	251	362	3.3e-38	-0.08	1.00		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
404	2gli	A	126	249	1.2e-31	-0.06	0.41		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
404	2gli	A	177	305	3.4e-33	0.43	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									CHAIN: C, D;	FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
404	2gli	A	197	335	3.4e-33			98.48	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
404	2gli	A	205	361	6.8e-30	0.21	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
405	1a68		51	135	1.4e-27	0.31	0.76		POTASSIUM CHANNEL KV1.1; CHAIN: NULL;	POTASSIUM CHANNELS, POTASSIUM CHANNELS, TETRAMERIZATION DOMAIN, X-RAY 2 STRUCTURE, APLYSIA KV1.1
405	1dsx	A	51	135	3.4e-26	0.22	0.92		KV1.2 VOLTAGE-GATED POTASSIUM CHANNEL; CHAIN: A, B, C, D, E, F, G, H;	SIGNALING PROTEIN VOLTAGE-GATED POTASSIUM CHANNEL, ASSEMBLY DOMAIN, TETRAMER
405	1exb	E	50	136	3.4e-27	0.27	0.98		KV BETA2 PROTEIN; CHAIN: A; POTASSIUM CHANNEL KV1.1; CHAIN: E;	METAL TRANSPORT ION CHANNEL, OXIDOREDUCTASE, BETA SUBUNIT
405	1qdv	A	51	149	3.4e-29	0.34	0.98		KV1.2 VOLTAGE-GATED POTASSIUM CHANNEL; CHAIN: A, B, C, D;	SIGNALING PROTEIN VOLTAGE-GATED POTASSIUM CHANNEL, TETRAMERIZATION DOMAIN, 2 INTRACELLULAR GATE, TETRAMER
405	1t1d	A	51	150	1.7e-29	0.16	0.98		POTASSIUM CHANNEL KV1.1; CHAIN: A;	PROTON TRANSPORT POTASSIUM CHANNELS, TETRAMERIZATION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
405	3kvt		50	151	1.7e-33	0.55	0.52		POTASSIUM CHANNEL PROTEIN SHAW; CHAIN: NULL;	POTASSIUM CHANNEL POTASSIUM CHANNEL, TETRAMERIZATION DOMAIN, MOLECULAR 2 RECOGNITION, ZINC-BINDING
407	1cdy		50	138	9.9e-12	0.36	0.77		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: NULL;	T-CELL SURFACE GLYCOPROTEIN IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC, LIPOPROTEIN, T-CELL SURFACE GLYCOPROTEIN
407	1cvs	C	50	145	1.3e-08	0.33	0.36		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR PGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
407	1cvs	D	50	145	2e-08	0.37	0.30		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR PGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
407	1dga	R	47	136	2e-08	0.26	1.00		POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PVR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS-RECEPTOR COMPLEX, VIRUS/VIRAL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
407	1evt	C	47	136	2e-08	0.46	0.57		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	PROTEIN, RECEPTOR GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
407	1fhg	A	47	138	6.6e-08	0.19	1.00		TELOKIN; CHAIN: A	CONTRACTILE PROTEIN IMMUNOGLOBULIN FOLD, BETA BARREL
407	1nct		43	136	6.8e-09	0.53	0.99		TTIN; CHAIN: NULL;	MUSCLE PROTEIN CONNECTIN, NEXTM5; CELL ADHESION, GLYCOPROTEIN, TRANSMEMBRANE, REPEAT, BRAIN, 2 IMMUNOGLOBULIN FOLD, ALTERNATIVE SPLICING, SIGNAL, 3 MUSCLE PROTEIN
407	1trm		43	136	6.8e-09	0.92	0.96		MUSCLE PROTEIN TTIN MODULE M5 (CONNECTIN) ITNM 3 (NMR, MINIMIZED AVERAGE STRUCTURE) ITNM 4 ITNM 58	
407	1wio	A	49	138	6.6e-09	-0.07	0.46		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC LIPOPROTEIN, POLYMORPHISM
408	1c3g	A	140	310	3.3e-53	0.34	1.00		HEAT SHOCK PROTEIN 40; CHAIN: A;	CHAPERONE BETA SHEETS SHORT HELICES
408	1c3g	A	145	310	6.8e-33	0.26	1.00		HEAT SHOCK PROTEIN 40; CHAIN: A;	CHAPERONE BETA SHEETS SHORT HELICES
408	1hdj		1	75	2.3e-29	0.59	1.00		HUMAN HSP40; CHAIN:	MOLECULAR CHAPERONE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									NULL;	HDJ-1; MOLECULAR CHAPERONE
408	1hdj		1	75	2.3e-29			93.05	HUMAN HSP40; CHAIN: NULL;	MOLECULAR CHAPERONE HDJ-1; MOLECULAR CHAPERONE
408	1hdj		1	76	3.4e-25	0.75	1.00		HUMAN HSP40; CHAIN: NULL;	MOLECULAR CHAPERONE HDJ-1; MOLECULAR CHAPERONE
410	1b8t	A	274	469	6.6e-15			63.60	CRP1; CHAIN: A;	CONTRACTILE LIM DOMAIN, CRP, NMR, MUSCLE DIFFERENTIATION, CONTRACTILE
410	1mey	C	110	191	1e-45	0.19	0.96		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
410	1mey	C	138	219	3.4e-46	0.34	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
410	1mey	C	166	247	1.2e-47	0.78	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
410	1mey	C	194	275	5.1e-48	0.10	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
410	1mey	C	222	303	1.7e-49	0.06	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
410	1mey	C	250	331	1.7e-49	-0.19	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
410	1mey	C	278	359	1.7e-50	0.15	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
410	1mey	C	306	387	1.7e-50	0.13	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
410	1mey	C	334	415	6.8e-51	0.41	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
410	1mey	C	334	416	6.8e-51			95.94	DNA; CHAIN: A, B, D, E;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID No:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
410	1mey	C	362	443	1.2e-50	0.26	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
410	1mey	C	390	471	3.4e-50	0.51	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
410	1mey	C	418	475	1.5e-33	0.26	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
410	1mey	C	54	135	6.8e-43	-0.32	0.13		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
410	1mey	C	82	163	1.7e-44	-0.41	0.05		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										STRUCTURE, COMPLEX (ZINC FINGER/DNA)
410	1tf6	A	111	256	5.1e-36	0.08	0.66		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
410	1tf6	A	223	368	1e-38	-0.09	0.83		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
410	1tf6	A	306	474	1.7e-36			104.78	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
410	1tf6	A	335	473	1.7e-36	0.01	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
410	1tf6	A	55	205	3.4e-34	-0.46	0.43		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
410	1ubd	C	120	220	2e-38	0.20	1.00		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
410	1ubd	C	139	248	3.3e-50			88.64	YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
410	1ubd	C	141	248	3.3e-50	0.02	1.00		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
410	1ubd	C	193	304	6.6e-56	0.09	1.00		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
410	1ubd	C	202	303	1e-34	-0.02	1.00		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
410	1ubd	C	224	331	1.2e-35	-0.12	1.00		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
410	1ubd	C	249	359	3.3e-53	-0.07	1.00		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
410	1ubd	C	276	415	2e-53	0.02	0.70		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
410	1ubd	C	286	387	1.2e-35	-0.11	1.00		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
410	1ubd	C	332	443	3.3e-54	0.12	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
410	1ubd	C	339	443	8.5e-36	0.29	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
410	1ubd	C	360	470	2e-51	0.29	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
410	1ubd	C	370	471	1e-34	0.20	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
410	1ubd	C	62	163	3.4e-30	-0.26	0.33		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
410	1ubd	C	90	191	5.1e-32	-0.24	0.72		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
410	2gli	A	174	305	3.4e-34	0.49	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
410	2gli	A	258	386	3.4e-34	0.16	0.99		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
410	2gli	A	26	162	5.1e-31	-0.34	0.11		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
410	2gli	A	334	473	1.7e-34			94.24	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER GLI; GLI, ZINC

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
410	2gli	A	342	470	1.7e-34	0.45	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
410	2gli	A	82	218	1e-32	0.02	0.90		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
411	1a03	A	1	92	1.4e-18			70.10	CALCYCLIN (RABBIT, CA2+); CHAIN: A, B;	CALCIUM-BINDING PROTEIN 2A9, CACY, S100A6, PRA; CALCIUM-BINDING PROTEIN, EF-HAND, S-100 PROTEIN, NMR
411	1a4p	A	2	96	3.3e-21			87.95	S100A10; CHAIN: A, B;	CALCIUM/PHOSPHOLIPID BINDING PROTEIN P11, CALPACTIN LIGHT CHAIN; S100 FAMILY, EF-HAND PROTEIN, LIGAND OF ANNEXIN II, 2
411	1b4c	A	2	93	6.8e-22			84.29	S-100 PROTEIN, BETA CHAIN; CHAIN: A, B;	CALCIUM/PHOSPHOLIPID BINDING PROTEIN METAL BINDING PROTEIN S100B, S100BETA; S100BETA, S100B, NMR, DIPOLAR COUPLINGS, EF-HAND, S100 2 PROTEIN, CALCIUM-BINDING PROTEIN, FOUR-HELIX BUNDLE, THREE- 3 DIMENSIONAL STRUCTURE SOLUTION STRUCTURE
411	1blq		4	84	1.5e-24	0.02	-0.07		N-TROPONIN C; CHAIN: NULL;	CALCIUM-BINDING PROTEIN S100C; CALCIUM-BINDING, REGULATION, FU

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
411	1exr	A	4	99	1.4e-31	0.00	-0.15		CALMODULIN; CHAIN: A;	TROPONIN C, SKELETAL MUSCLE, 2 CONTRACTION
411	1mho		3	90	1.7e-20			83.93	S-100 PROTEIN; CHAIN: NULL;	METAL TRANSPORT CALMODULIN, HIGH RESOLUTION, DISORDER
411	1top		4	99	6.8e-27	0.21	0.07		CONTRACTILE SYSTEM PROTEIN TROPONIN C 1TOP 3	CALCIUM-BINDING CALCIUM-BINDING, ZINC, METAL-BINDING, ACETYLATION
411	1trf		4	84	1.5e-24	0.26	0.15		MUSCLE PROTEIN TROPONIN C (TRIC FRAGMENT) (APO FORM) (NMR, 1 STRUCTURE) 1TRF 3	
413	1bj8		15	118	3.4e-16			51.99	GP130; CHAIN: NULL;	RECEPTOR RECEPTOR, SIGNAL TRANSDUCER OF IL-6 TYPE CYTOKINES, THIRD 2 N-TERMINAL DOMAIN, TRANSMEMBRANE, GLYCOPROTEIN
413	1bj8		26	111	3.4e-16	0.28	0.81		GP130; CHAIN: NULL;	RECEPTOR RECEPTOR, SIGNAL TRANSDUCER OF IL-6 TYPE CYTOKINES, THIRD 2 N-TERMINAL DOMAIN, TRANSMEMBRANE, GLYCOPROTEIN
413	1bp3	B	21	218	1.2e-21			57.27	GROWTH HORMONE; CHAIN: A; PROLACTIN RECEPTOR; CHAIN: B;	HORMONE/GROWTH FACTOR HORMONE, RECEPTOR, HORMONE/GROWTH FACTOR
413	1bpv		116	205	5.1e-09	0.00	-0.14		TITIN; CHAIN: NULL;	CONNECTIN A71, CONNECTIN, TITIN,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
413	1bpv		21	119	2.3e-18	0.12	0.28		TTTN; CHAIN: NULL;	CONNECTIN, FIBRONECTIN TYPE III
413	1bpv		32	114	1.7e-14	-0.10	0.28		TTTN; CHAIN: NULL;	CONNECTIN A71, CONNECTIN; TTTN, CONNECTIN, FIBRONECTIN TYPE III
413	1bqu	A	20	231	1.4e-16			52.35	GPI30; CHAIN: A, B;	SIGNALING PROTEIN CYTOKINE RECEPTOR, GLYCOPROTEIN 130, GPI30, INTERLEUKINE 6 2 RECEPTOR BETA SUBUNIT, SIGNALING PROTEIN
413	1bqu	A	26	124	1.2e-17	0.04	0.37		GPI30; CHAIN: A, B;	SIGNALING PROTEIN CYTOKINE RECEPTOR, GLYCOPROTEIN 130, GPI30, INTERLEUKINE 6 2 RECEPTOR BETA SUBUNIT, SIGNALING PROTEIN
413	1c8p	A	18	114	1.3e-14	0.09	0.15		CYTOKINE RECEPTOR COMMON BETA CHAIN; CHAIN: A;	MEMBRANE PROTEIN BETA SANDWICH, CYTOKINE RECEPTOR, FN3 DOMAIN
413	1cfb		18	224	5.1e-24			64.96	NEURAL ADHESION MOLECULE DROSOPHILA NEUROGLIAN (CHYMOTRYPTIC FRAGMENT CONTAINING THE 1CFB 3 TWO AMINO PROXIMAL FIBRONECTIN TYPE III REPEATS 1CFB 4 (RESIDUES 610 - 814)) 1CFB 5	
413	1cto		21	108	2e-12	-0.31	0.24		GRANULOCYTE COLONY-STIMULATING FACTOR RECEPTOR; CHAIN: NULL;	BINDING PROTEIN BINDING PROTEIN, CYTOKINE RECEPTOR
413	1eer	B	18	114	9.9e-17	-0.05	0.58		ERYTHROPOIETIN;	COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									CHAIN: A; ERYTHROPOIETIN RECEPTOR; CHAIN: B, C; RECEPTOR, SIGNAL 2 TRANSDUCTION, HEMATOPOIETIC CYTOKINE, CYTOKINE RECEPTOR 3 CLASS 1, COMPLEX (CYTOKINE/RECEPTOR)	(CYTOKINE/RECEPTOR) EPOBP; ERYTHROPOIETIN, ERYTHROPOIETIN RECEPTOR, SIGNAL 2 TRANSDUCTION, HEMATOPOIETIC CYTOKINE, CYTOKINE RECEPTOR 3 CLASS 1, COMPLEX (CYTOKINE/RECEPTOR)
413	1fna		28	107	8.5e-14	-0.27	0.69		CELL ADHESION PROTEIN FIBRONECTIN CELL- ADHESION MODULE TYPE III-10 IFNA 3	
413	1fnf		21	362	8.5e-32			92.46	FIBRONECTIN; 1FNF 6 CHAIN: NULL; 1FNF 7	CELL ADHESION PROTEIN RGD, EXTRACELLULAR MATRIX 1FNF 18
413	1fnh	A	23	293	1e-27			81.90	FIBRONECTIN; CHAIN: A;	HEPARIN AND INTEGRIN BINDING HEPARIN AND INTEGRIN BINDING
413	1fnh	A	3	106	5.1e-15	-0.65	0.19		FIBRONECTIN; CHAIN: A;	HEPARIN AND INTEGRIN BINDING HEPARIN AND INTEGRIN BINDING
413	1mfn		21	207	1.7e-26			60.33	FIBRONECTIN; CHAIN: NULL;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN, RGD, EXTRACELLULAR MATRIX, 2 HEPARIN- BINDING, GLYCOPROTEIN
413	1mfn		23	157	9.9e-19	-0.02	0.07		FIBRONECTIN; CHAIN: NULL;	CELL ADHESION PROTEIN, RGD, EXTRACELLULAR MATRIX, 2 HEPARIN- BINDING, GLYCOPROTEIN
413	1mfn		4	107	1.7e-14	-0.40	0.64		FIBRONECTIN; CHAIN: NULL;	CELL ADHESION PROTEIN, RGD, EXTRACELLULAR MATRIX, 2 HEPARIN- BINDING, GLYCOPROTEIN
413	1gg3	A	21	231	1.2e-25			66.57	INTEGRIN BETA-4	STRUCTURAL PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									SUBUNIT; CHAIN: A, B;	INTEGRIN, HEMIDESMOSOME, FIBRONECTIN, CARCINOMA, STRUCTURAL 2 PROTEIN
413	1qg3	A	26	208	1.2e-25	0.16	0.35		INTEGRIN BETA-4 SUBUNIT; CHAIN: A, B;	STRUCTURAL PROTEIN INTEGRIN, HEMIDESMOSOME, FIBRONECTIN, CARCINOMA, STRUCTURAL 2 PROTEIN
413	1qr4	A	23	206	3.4e-18			70.56	TENASCIN; CHAIN: A, B;	STRUCTURAL PROTEIN TENASCIN, FIBRONECTIN TYPE-III, HEPARIN, EXTRACELLULAR 2 MATRIX, ADHESION, FUSION PROTEIN, STRUCTURAL PROTEIN
413	1qr4	A	27	204	3.4e-18	0.00	0.28		TENASCIN; CHAIN: A, B;	STRUCTURAL PROTEIN TENASCIN, FIBRONECTIN TYPE-III, HEPARIN, EXTRACELLULAR 2 MATRIX, ADHESION, FUSION PROTEIN, STRUCTURAL PROTEIN
413	1tff		26	107	1.7e-14	-0.29	0.37		GLYCOPROTEIN FIBRONECTIN (TENTH TYPE III MODULE) (NMR, 36 STRUCTURES) 1TTF 3 FIBRONECTIN; CHAIN: A;	
413	2fmb	A	21	110	3e-15	0.29	0.95			PROTEIN BINDING ED-B, FIBRONECTIN, TYPEIII DOMAIN, ANGIOGENESIS, PROTEIN 2 BINDING
414	1alh	A	122	195	1.4e-29	-0.68	0.04		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
414	1alh	A	143	229	5.1e-29	-0.42	0.01		C; QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
414	1alh	A	531	595	9.9e-06	0.64	0.25		C; QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
414	1bbo		531	120	0.0033	0.08	0.46		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11 1BBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) 1BBO 4	
414	1mey	C	119	195	6.8e-47	-0.63	0.21		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
414	1mey	C	2	64	5.1e-38	-0.18	0.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
414	1mey	C	531	604	0.0012	0.37	0.41		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
414	1mey	G	168	195	3.4e-13	-0.33	0.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	STRUCTURE, COMPLEX (ZINC FINGER/DNA)
414	1mey	G	37	64	6.8e-14	-0.19	0.35		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
414	1ubd	C	122	229	5.1e-34	-0.49	0.05		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
414	1ubd	C	43	193	3e-16	-0.26	0.81		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
414	1ubd	C	527	600	1.6e-07	0.21	0.36		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
414	2drp	A	531	595	3.3e-06	0.48	0.03		COMPLEX(TRANSCRIPTIO N REGULATION/DNA) TRAMTRACK PROTEIN (TWO ZINC-FINGER PEPTIDE) COMPLEXED WITH 2DRP 3 DNA 2DRP 4	FINGER PROTEIN, DNA- PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
415	1chc		12	61	9.9e-14	-0.12	0.66		VIRUS EQUINE HERPES VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	
415	1fbv	A	16	56	1e-06	-0.52	0.48		SIGNAL TRANSDUCTION PROTEIN CBL; CHAIN: A; ZAP-70 PEPTIDE; CHAIN: B; UBIQUITIN- CONJUGATING ENZYME B12-18 KDA UBC7; CHAIN: C;	LIGASE CBL, UBC7, ZAP- 70, E2, UBIQUITIN, E3, PHOSPHORYLATION, 2 TYROSINE KINASE, UBIQUITINATION, PROTEIN DEGRADATION,
415	1g25	A	12	68	3.3e-13	0.50	0.40		CDK-ACTIVATING KINASE ASSEMBLY FACTOR MAT1; CHAIN: A;	METAL BINDING PROTEIN RING FINGER PROTEIN MAT1; RING FINGER (C3HC4)
415	1rmd		3	104	5.1e-12	-0.36	0.05		RAG1; CHAIN: NULL;	DNA-BINDING PROTEIN V(D)J RECOMBINATION ACTIVATING PROTEIN 1; RAG1, V(D)J RECOMBINATION, ANTIBODY, MAD, RING FINGER, 2 ZINC BINUCLEAR CLUSTER, ZINC FINGER, DNA-BINDING PROTEIN
416	1a06		87	398	5.1e-77			110.80	CALCIUM/CALMODULIN- DEPENDENT PROTEIN KINASE; CHAIN: NULL;	KINASE KINASE, SIGNAL TRANSDUCTION, CALCIUM/CALMODULIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
416	1a06		97	384	5.1e-77	-0.24	1.00		CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE; CHAIN: NULL;	KINASE KINASE, SIGNAL TRANSDUCTION, CALCIUM/CALMODULIN
416	1a60		66	400	1.2e-35			105.92	PROTEIN KINASE CK2/ALPHA-SUBUNIT; CHAIN: NULL;	TRANSFERASE TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, CASEIN KINASE, 2 SER/THR KINASE
416	1apm	E	61	413	0	0.11	0.96		TRANSFERASE(PHOSPHO TRANSFERASE) \$C-/AMP\$-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (\$C/APK\$) 1APM 3 (CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (S139A\$) COMPLEX WITH THE PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 1APM 6	
416	1apm	E	68	410	0			178.54	TRANSFERASE(PHOSPHO TRANSFERASE) \$C-/AMP\$-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (\$C/APK\$) 1APM 3 (CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (S139A\$) COMPLEX WITH THE PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 1APM 6	
416	1aq1		94	407	1.4e-51			98.38	CYCLIN-DEPENDENT PROTEIN KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; PROTEIN KINASE, CELL CYCLE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
416	1b6c	B	65	382	2e-25			93.15	FK506-BINDING PROTEIN; CHAIN: A, C, E, G; TGF-B SUPERFAMILY RECEPTOR TYPE I; CHAIN: B, D, F, H;	COMPLEX (ISOMERASE/PROTEIN KINASE) FKBP12; SERINE/THREONINE-PROTEIN KINASE RECEPTOR R4; COMPLEX (ISOMERASE/PROTEIN KINASE), RECEPTOR 2 SERINE/THREONINE KINASE
416	1blx	A	72	400	6.8e-45			94.82	CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
416	1cmk	E	60	410	0			180.11	PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT 1CMK 3 (E.C.2.7.1.37) 1CMK 4	
416	1cmk	E	61	413	0	0.04	0.94		PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT 1CMK 3 (E.C.2.7.1.37) 1CMK 4	
416	1ctp	E	60	407	0			179.90	TRANSFERASE(PHOSPHOTRANSFERASE) CAMP-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (CAPK) 1CTP 3	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
416	1ctp	E	61	381	0	0.24	1.00		(CATALYTIC SUBUNIT) ICTP 4 TRANSFERASE(PHOSPHO TRANSFERASE) CAMP-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (CAPK) ICTP 3 (CATALYTIC SUBUNIT) ICTP 4	
416	1ez3	A	398	515	9.9e-13	0.55	0.18		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
416	1f3m	C	84	316	3.3e-63	-0.06	1.00		SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: A, B; SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: C, D;	TRANSFERASE KINASE DOMAIN, AUTOINHIBITORY FRAGMENT, HOMODIMER
416	1hcl		94	407	1e-53			109.49	HUMAN CYCLIN-DEPENDENT KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; TRANSFERASE, SERINE/THREONINE PROTEIN KINASE, ATP-BINDING, 2 CELL CYCLE, CELL DIVISION, MITOSIS, PHOSPHORYLATION
416	ljnk		78	446	5.1e-39			91.45	C-JUN N-TERMINAL KINASE; CHAIN: NULL;	TRANSFERASE JNK3; TRANSFERASE, JNK3 MAP KINASE,
416	lkoa		75	357	1.2e-71	0.13	1.00		TWITCHIN; CHAIN: NULL;	SERINE/THREONINE PROTEIN 2 KINASE KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
416	lkob	A	71	382	5.1e-77	0.21	1.00		TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
416	lkob	A	73	422	5.1e-77			142.88	TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
416	lp38		72	439	3.4e-44			105.18	MAP KINASE P38; CHAIN:	TRANSFERASE MITOGEN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									NULL;	ACTIVATED PROTEIN KINASE; TRANSFERASE, MAP KINASE, SERINE/THREONINE-PROTEIN KINASE, 2 P38
416	1phk		93	367	1.7e-80			126.81	PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE-PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING
416	1phk		95	356	1.7e-80	0.32	1.00		PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE-PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING
416	1pme		85	431	1.2e-39			108.39	ERK2; CHAIN: NULL;	TRANSFERASE MAP KINASE, SERINE/THREONINE PROTEIN KINASE, TRANSFERASE
416	1tki	A	91	361	1e-57	0.11	1.00		TTIN; CHAIN: A, B;	SERINE KINASE SERINE KINASE, TITIN, MUSCLE, AUTOINHIBITION
416	1tki	A	91	405	6.6e-64			119.19	TTIN; CHAIN: A, B;	SERINE KINASE SERINE KINASE, TITIN, MUSCLE, AUTOINHIBITION
416	1tki	A	95	317	6.6e-64	0.07	1.00		TTIN; CHAIN: A, B;	SERINE KINASE SERINE KINASE, TITIN, MUSCLE, AUTOINHIBITION
416	3erk		62	436	5.1e-42			108.88	EXTRACELLULAR REGULATED KINASE 2; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE, MAP 2, ERK2; TRANSFERASE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										SERINE/THREONINE-PROTEIN KINASE, MAP KINASE, 2 ERK2
417	1a0j	A	18	241	0			172.43	TRYPSIN; CHAIN: A, B, C, D;	SERINE PROTEASE SERINE PROTEINASE, TRYPSIN, HYDROLASE
417	1a0j	A	19	241	0	0.76	1.00		TRYPSIN; CHAIN: A, B, C, D;	SERINE PROTEASE SERINE PROTEINASE, TRYPSIN, HYDROLASE
417	1a0l	A	18	241	3.4e-81			124.61	BETA-TRYPTASE; CHAIN: A, B, C, D;	SERINE PROTEINASE TRYPSIN-LIKE SERINE PROTEINASE, TETRAMER, HEPARIN, ALLERGY, 2 ASTHMA
417	1aks	A	19	142	1.7e-49	0.26	1.00		ALPHA TRYPSIN; CHAIN: A, B;	SERINE PROTEASE HYDROLASE, SERINE PROTEASE
417	1aks	B	143	241	1.4e-47	-0.48	0.99		ALPHA TRYPSIN; CHAIN: A, B;	SERINE PROTEASE HYDROLASE, SERINE PROTEASE
417	1ao5	A	14	241	1.7e-79			132.64	GLANDULAR KALLIKREIN-13; CHAIN: A, B;	SERINE PROTEASE PRORENIN CONVERTING ENZYME (PRECE), EPIDERMAL GLANDULAR KALLIKREIN, SERINE PROTEASE, PROTEIN MATURATION
417	1aut	C	18	239	5.1e-72			124.65	ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR), AUTO-PROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE, PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
417	1bbx	B	20	144	3.4e-35	-0.05	0.87		ALPHA THROMBIN; CHAIN: A, B, F, E;	SERINE PROTEASE SERINE PROTEASE HEADER

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
417	1bix	F	149	238	1.7e-32	-0.62	0.55		ALPHA THROMBIN; CHAIN: A, B, F, E;	HETNAM SERINE PROTEASE SERINE PROTEASE HEADER HETNAM
417	1bio		17	239	1.5e-68			133.21	COMPLEMENT FACTOR D; CHAIN: NULL;	SERINE PROTEASE SERINE PROTEASE, HYDROLASE, COMPLEMENT, FACTOR D, CATALYTIC 2 TRIAD, SELF-REGULATION
417	1bgj	A	17	241	6.8e-87			154.29	PLASMINOGEN ACTIVATOR; CHAIN: A, B; GLU-GLY-ARG-CHLOROMETHYLKETONE INHIBITOR; CHAIN: E, F;	BLOOD CLOTTING TSV-PA; FIBRINOLYSIS, PLASMINOGEN ACTIVATOR, SERINE PROTEINASE, 2 SNAKE VENOM, COMPLEX (HYDROLASE/INHIBITOR), BLOOD CLOTTING
417	1bgj	A	18	241	6.8e-87	0.69	1.00		PLASMINOGEN ACTIVATOR; CHAIN: A, B; GLU-GLY-ARG-CHLOROMETHYLKETONE INHIBITOR; CHAIN: E, F;	BLOOD CLOTTING TSV-PA; FIBRINOLYSIS, PLASMINOGEN ACTIVATOR, SERINE PROTEINASE, 2 SNAKE VENOM, COMPLEX (HYDROLASE/INHIBITOR), BLOOD CLOTTING
417	1cgh	A	17	240	8.5e-71			126.65	CATHEPSIN G; CHAIN: A; PHOSPHONATE INHIBITOR SUC-VAL-PRO-PHEP-(OPH)2; CHAIN: S;	COMPLEX (SERINE PROTEASE/INHIBITOR) INFLAMMATION, INHIBITOR, SPECIFICITY, SERINE PROTEASE, 2 COMPLEX (SERINE PROTEASE/INHIBITOR)
417	1dpo		18	241	5.1e-96			165.05	TRYPSIN; CHAIN: NULL;	SERINE PROTEASE HYDROLASE, SERINE PROTEASE, DIGESTION, PANCREAS, ZYMOGEN, 2 SIGNAL, MULTIGENE FAMILY
417	1ekb	B	16	239	1e-79			136.64	ENTEROPEPTIDASE;	HYDROLASE/HYDROLASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
417	1ept	A	19	59	5.1e-15	-0.81	0.65		HYDROLASE (SERINE PROTEASE) PORCINE E-TRYPSIN (E.C.3.4.21.4) IEPT 3	INHIBITOR ENTEROKINASE, HEAVY CHAIN; ENTEROKINASE, LIGHT CHAIN; ENTEROPEPTIDASE, TRYPsinogen ACTIVATION, 2 HYDROLASE/HYDROLASE INHIBITOR
417	1fxy	A	18	241	1.7e-88			155.00	COAGULATION FACTOR XA-TRYPSIN CHIMERA; CHAIN: A; D-PHE-PRO-ARG-CHLOROMETHYLKETONE (PPACK) WITH CHAIN: I;	COMPLEX (PROTEASE/INHIBITOR) TRYPSIN, COAGULATION FACTOR XA, CHIMERA, PROTEASE, PPACK, 2 CHLOROMETHYLKETONE, COMPLEX (PROTEASE/INHIBITOR)
417	1fxy	A	19	241	1.7e-88	0.53	1.00		COAGULATION FACTOR XA-TRYPSIN CHIMERA; CHAIN: A; D-PHE-PRO-ARG-CHLOROMETHYLKETONE (PPACK) WITH CHAIN: I;	COMPLEX (PROTEASE/INHIBITOR) TRYPSIN, COAGULATION FACTOR XA, CHIMERA, PROTEASE, PPACK, 2 CHLOROMETHYLKETONE, COMPLEX (PROTEASE/INHIBITOR)
417	1gct	A	7	241	1.5e-78			125.17	HYDROLASE (SERINE PROTEINASE) GAMMA-*CHYMOTRYPSIN *A (E.C.3.4.21.1) (\$P*H 7.0) 1GCT 3	
417	1gg6	B	19	143	6.8e-42	0.27	0.94		GAMMA CHYMOTRYPSIN; CHAIN: A; GAMMA CHYMOTRYPSIN; CHAIN: B; GAMMA CHYMOTRYPSIN; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR CHYMOTRYPSIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
417	1klt		17	241	6.8e-72			125.55	CHYMASE; CHAIN: NULL;	SERINE PROTEASE SERINE PROTEASE, HYDROLASE, MAST CELL, ANGIOTENSIN, ALPHA 2 TOLUENESULFONIC ACID
417	1mct	A	18	241	0			173.18	COMPLEX(TRYPSIN INHIBITOR) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH INHIBITOR FROM BITTER 1MCT 3 GOURD 1MCT 4	
417	1mct	A	19	241	0	0.76	1.00		COMPLEX(TRYPSIN INHIBITOR) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH INHIBITOR FROM BITTER 1MCT 3 GOURD 1MCT 4	
417	1npm	A	16	239	5.1e-84			139.50	NEUROPSIN; CHAIN: A, B;	SERINE PROTEINASE SERINE PROTEINASE, GLYCOPROTEIN
417	1pfx	C	18	241	3.4e-77			115.05	FACTOR IXA; CHAIN: C, L.; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
417	1qrz	A	7	241	1e-81			114.05	PLASMINOGEN; CHAIN: A, B, C, D;	HYDROLASE MICROPLASMINOGEN, SERINE PROTEASE, ZYMOGEN, CHYMOTRYPSIN 2 FAMILY HYDROLASE
417	1sgf	A	25	241	6.8e-71			119.87	NERVE GROWTH FACTOR; CHAIN: A, B, G, X, Y, Z;	GROWTH FACTOR 7S NGF, GROWTH FACTOR (BET A - TU NGF), HYDROLASE - SERINE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										PROTEINASE 2 (GAMMA-NGF), INACTIVE SERINE PROTEINASE (ALPHA-NGF)
417	1sgf	G	16	241	1.7e-88			134.19	NERVE GROWTH FACTOR; CHAIN: A, B, G, X, Y, Z;	GROWTH FACTOR 7S NGF; GROWTH FACTOR (BETA-NGF), HYDROLASE - SERINE PROTEINASE 2 (GAMMA-NGF), INACTIVE SERINE PROTEINASE (ALPHA-NGF)
417	1sgf	G	22	241	1.7e-88	0.60	1.00		NERVE GROWTH FACTOR; CHAIN: A, B, G, X, Y, Z;	GROWTH FACTOR 7S NGF; GROWTH FACTOR (BETA-NGF), HYDROLASE - SERINE PROTEINASE 2 (GAMMA-NGF), INACTIVE SERINE PROTEINASE (ALPHA-NGF)
417	1slw	B	18	241	1.2e-97			165.58	ECOTIN; CHAIN: A; ANIONIC TRYPSIN; CHAIN: B;	COMPLEX (SERINE PROTEASE/INHIBITOR); TRYPSIN INHIBITOR; SERINE PROTEASE, INHIBITOR, COMPLEX, METAL BINDING SITES, 2 PROTEIN ENGINEERING, PROTEASE-SUBSTRATE INTERACTIONS, 3 METALLOPROTEINS
417	1slw	B	19	241	1.2e-97	0.67	1.00		ECOTIN; CHAIN: A; ANIONIC TRYPSIN; CHAIN: B;	COMPLEX (SERINE PROTEASE/INHIBITOR); TRYPSIN INHIBITOR; SERINE PROTEASE, INHIBITOR, COMPLEX, METAL BINDING SITES, 2 PROTEIN ENGINEERING, PROTEASE-SUBSTRATE INTERACTIONS, 3 METALLOPROTEINS
417	1ton		18	241	1e-79			145.44	HYDROLASE(SERINE PROTEINASE) TONIN (E.C. NUMBER NOT ASSIGNED) ITON 4	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
417	ltn	A	18	241	1.7e-100			177.94	HYDROLASE (SERINE PROTEINASE) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH THE INHIBITOR ITRN 3 DIISOPROPYL-FLUOROPHOSPHOFLUORIDE (DFP) ITRN 4 HUMAN TRYPSIN, DFP INHIBITED ITRN 6	
417	ltn	A	19	241	1.7e-100	0.69	1.00		HYDROLASE (SERINE PROTEINASE) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH THE INHIBITOR ITRN 3 DIISOPROPYL-FLUOROPHOSPHOFLUORIDE (DFP) ITRN 4 HUMAN TRYPSIN, DFP INHIBITED ITRN 6	
417	lucy	E	149	238	5.1e-32	-0.44	0.58		THROMBIN; CHAIN: L, H, E, I, K, M, N; FIBRINOPEPTIDE A-ALPHA; CHAIN: F, G, I;	COMPLEX (SERINE PROTEASE/COAGULATION) COMPLEX (SERINE PROTEASE/COAGULATION), SERINE, PROTEASE, 2 THROMBIN
417	lucy	H	20	144	3.4e-34	0.11	0.98		THROMBIN; CHAIN: L, H, E, I, K, M, N; FIBRINOPEPTIDE A-ALPHA; CHAIN: F, G, I;	COMPLEX (SERINE PROTEASE/COAGULATION) COMPLEX (SERINE PROTEASE/COAGULATION), SERINE, PROTEASE, 2 THROMBIN
417	lycp	M	149	238	5.1e-32	-0.47	0.59		ALPHA THROMBIN; CHAIN: L, H; EPSILON THROMBIN; CHAIN: J, K, M; FIBRINOPEPTIDE A-ALPHA; CHAIN: F, N;	COMPLEX (SERINE PROTEASE/PEPTIDE) FIBRINOPEPTIDE-A, COMPLEX (SERINE PROTEASE/PEPTIDE), 2 THROMBIN
417	2hnt	C	20	70	6.8e-12	-0.39	0.05		SERINE PROTEASE GAMMA-THROMBIN 2HNT	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
417	2pka	A	21	94	3.4e-25	0.05	0.98		SERINE PROTEINASE KALLIKREIN A (E.C.3.4.21.8) 2PKA 4	
417	2pka	B	94	241	3.4e-54	-0.25	1.00		SERINE PROTEINASE KALLIKREIN A (E.C.3.4.21.8) 2PKA 4	
417	2tbs		18	241	5.1e-97			168.64	HYDROLASE/SERINE PROTEINASE TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH BENZAMIDINE INHIBITOR 2TBS 3	
417	2tbs		19	240	5.1e-97	0.65	1.00		HYDROLASE/SERINE PROTEINASE TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH BENZAMIDINE INHIBITOR 2TBS 3	
417	3rp2	A	17	239	1.4e-69			116.86	SERINE PROTEINASE RAT MAST CELL PROTEASE J15 (RMCP15) 3RP2 4	
417	5ptp		18	241	8.5e-98			170.27	BETA TRYPSIN; CHAIN: NULL;	SERINE PROTEASE HYDROLASE, SERINE PROTEASE, DIGESTION, PANCREAS, 2 ZYMOGEN, SIGNAL
417	5ptp		29	241	8.5e-98	0.84	1.00		BETA TRYPSIN; CHAIN: NULL;	SERINE PROTEASE HYDROLASE, SERINE PROTEASE, DIGESTION, PANCREAS, 2 ZYMOGEN, SIGNAL
422	1bql	H	34	197	3.4e-06	-0.29	0.12		COMPLEX (ANTIBODY/ANTIGEN) HYHEL-5 FAB COMPLEXED WITH BOBWHITE QUAIL LYSOZYME 1BQL 3 1BQL 95	
422	1cle	H	23	197	3.4e-05	-0.18	0.01		CATALYTIC ANTIBODY IE9 (LIGHT CHAIN);	IMMUNE SYSTEM CATALYTIC ANTIBODY, R1

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
422	1cic	D	21	228	1.4e-05			65.86	CHAIN: L; CATALYTIC ANTIBODY I19 (HEAVY CHAIN); CHAIN: H; IG HEAVY CHAIN V REGIONS; CHAIN: A; IG HEAVY CHAIN V REGIONS; CHAIN: B; IG HEAVY CHAIN V REGIONS; CHAIN: C; IG HEAVY CHAIN V REGIONS; CHAIN: D; IG HEAVY CHAIN V	DIELS-ALDER, IMMUNOGLOBULIN
422	1cic	D	35	197	1.4e-05	-0.09	0.42		IMMUNOGLOBULIN V REGIONS; CHAIN: A; IG HEAVY CHAIN V REGIONS; CHAIN: B; IG HEAVY CHAIN V REGIONS; CHAIN: C; IG HEAVY CHAIN V REGIONS; CHAIN: D; IG HEAVY CHAIN V	IMMUNOGLOBULIN, FAB COMPLEX, IDIOTOPE, ANTI-IDIOTOPE
422	1fns	H	35	197	1.7e-05	0.17	0.01		IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: L; IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: H; VON WILLEBRAND FACTOR; CHAIN: A;	IMMUNE SYSTEM VON WILLEBRAND FACTOR, GLYCOPROTEIN IBA (A:ALPHA) BINDING, 2 COMPLEX (WILLEBRAND/IMMUNOGLOBULIN), BLOOD COAGULATION TYPE 3 2B VON WILLEBRAND DISEASE
422	1frg	H	23	226	1e-05			66.17	IMMUNOGLOBULIN/VRUS HEMAGGLUTININ IGG2A FAB FRAGMENT (FAB 26/9) COMPLEXED WITH INFLUENZA IFRG 3 HEMAGGLUTININ HA1 (STRAIN X47) (RESIDUES 101 - 108) IFRG 4	
422	1fvd	B	21	228	5.1e-05			66.74	IMMUNOGLOBULIN FAB FRAGMENT OF	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
422	1mlb	B	24	197	3.4e-07	-0.26	0.07		HUMANIZED ANTIBODY 4D5, VERSION 4 1FVD 3	
422	1nfd	F	34	197	1.7e-05	0.07	0.42		IMMUNOGLOBULIN FAB D44.1 (IGG1,KAPPA) (BALB/C MOUSE, MONOCLONAL ANTIBODY) 1MLB 5	
422	1nld	H	21	228	5.1e-07			67.89	N15 ALPHA-BETA T-CELL RECEPTOR; CHAIN: A, B, C, D; H57 FAB; CHAIN: E, F, G, H	COMPLEX (IMMUNORECEPTOR/IMMUNOGLOBULIN) COMPLEX (IMMUNORECEPTOR/IMMUNOGLOBULIN)
422	1nld	H	23	229	5.1e-07	0.02	0.19		FAB1583; CHAIN: L, H	IMMUNOGLOBULIN FAB FRAGMENT, IMMUNOGLOBULIN
422	1r24	B	34	151	3.4e-05	-0.34	0.13		FAB1583; CHAIN: L, H	IMMUNOGLOBULIN FAB FRAGMENT, IMMUNOGLOBULIN
422	8fab	B	21	226	0.00017			65.00	IGG3-KAPPA ANTIBODY (LIGHT CHAIN); CHAIN: A, C; IGG3-KAPPA ANTIBODY (HEAVY CHAIN); CHAIN: B, D; IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	IMMUNE SYSTEM PRELIMINARY, IMMUNE SYSTEM
423	1adq	L	733	863	6.8e-13	0.08	0.09		IGG4 REA; CHAIN: A; RF-AN IGM/LAMBDA; CHAIN: H, L;	COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN) COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN), RHEUMATOID FACTOR 2 AUTO-ANTIBODY COMPLEX
423	1bih	A	367	715	1.4e-23	0.12	-0.02		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
423	1b1h	A	556	888	1.7e-29	0.32	0.24		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
423	1b1j8		442	541	2e-11	0.20	0.04		GPI30; CHAIN: NULL;	RECEPTOR RECEPTOR, SIGNAL TRANSDUCER OF IL-6 TYPE CYTOKINES, THIRD 2 N-TERMINAL DOMAIN, TRANSMEMBRANE, GLYCOPROTEIN
423	1bpv		442	528	3e-12	-0.07	0.51		TTIN; CHAIN: NULL;	CONNECTIN A71, CONNECTIN; TITIN, CONNECTIN, FIBRONECTIN TYPE III
423	1cs6	A	300	628	1.7e-21	-0.14	0.04		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
423	1cs6	A	351	722	1.7e-34	-0.00	0.27		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
423	1cs6	A	441	814	5.1e-27	-0.03	0.13		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
423	1cs6	A	120	888	1.5e-31	0.06	0.33		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
423	1cvs	C	354	536	8.5e-20	-0.09	0.01		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGF, IMMUNOGLOBULIN LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
423	1cvs	C	540	725	8.5e-11	0.00	-0.11		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGF, IMMUNOGLOBULIN LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
423	1cvs	C	729	888	8.5e-26	0.05	0.36		FIBROBLAST GROWTH	GROWTH FACTOR/GROWTH FACTOR RECEPTOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN- LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
423	1cvs	D	629	813	1.7e-25	0.19	-0.07		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN- LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
423	1cvs	D	729	888	6.8e-25	0.26	0.70		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN- LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
423	1dbb	H	363	477	1.7e-06	-0.31	0.10		IMMUNOGLOBULIN FAB' FRAGMENT OF THE DB3 ANTI-STERIOD MONOCLONAL ANTIBODY 1DBB 3 (IGG1, SUBGROUP 2A, KAPPA 1) COMPLEX WITH PROGESTERONE 1DBB 4	
423	1epf	A	359	520	5.1e-17	0.19	0.09		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
423	1epf	A	550	709	1.7e-11	0.22	0.10		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
423	1epf	A	643	801	1.4e-13	0.02	-0.07		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B,	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
423	1epf	A	731	888	5.1e-20	0.11	0.19		C, D; NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	FOLD, GLYCOPROTEIN CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
423	1ev2	E	646	813	1.2e-22	-0.09	0.03		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
423	1ev2	G	646	817	1.7e-23	0.05	0.00		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
423	1ev2	G	739	888	1.4e-23	-0.00	0.27		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
423	1evt	C	354	536	5.1e-18	-0.07	0.21		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
423	1evt	C	540	725	1.7e-10	0.12	-0.09		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
423	1evt	C	729	888	8.5e-24	0.15	0.18		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFI; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
423	1f2q	A	742	888	3.3e-13	0.09	0.36		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RI-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
423	1f6a	A	667	816	6.6e-12	0.06	0.63		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IG EPSILON CHAIN C REGION; CHAIN: B, D;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
423	1f6a	A	737	888	6.6e-15	0.06	0.18		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IG EPSILON CHAIN C REGION; CHAIN: B, D;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
423	1fcg	A	667	802	9.9e-12	-0.00	-0.12		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32
423	1fhg	A	351	440	5.1e-16	0.31	0.84		TELOKIN; CHAIN: A	FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
423	1fhg	A	538	629	1e-09	0.39	0.75		TELOKIN; CHAIN: A	CONTRACTILE PROTEIN IMMUNOGLOBULIN FOLD, BETA BARREL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
423	1itb	B	574	813	3.3e-15	-0.22	0.09		INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	BETA BARREL COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
423	1itb	B	662	869	3.3e-24	0.07	0.27		INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
423	1itb	B	742	888	3.3e-15	0.19	0.69		INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
423	1koa		351	430	1.5e-12	-0.15	0.37		TWITCHIN; CHAIN: NULL;	KINASE KINASE, TWITCHIN; INTRASTERIC REGULATION
423	1mco	H	430	818	1.2e-11			83.10	IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (GG1) (MCG) WITH A HINGE DELETION IMCO 3	
423	1nct		353	429	6.8e-14	0.19	0.28		TITIN; CHAIN: NULL;	MUSCLE PROTEIN CONNECTIN, NEXTMS; CELL ADHESION, GLYCOPROTEIN, TRANSMEMBRANE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
423	1nct		626	722	6.6e-10	0.13	0.19		TITIN; CHAIN: NULL;	MUSCLE PROTEIN CONNECTIN, NEXTM5; CELL ADHESION, GLYCOPROTEIN, TRANSMEMBRANE, REPEAT, BRAIN, 2 IMMUNOGLOBULIN FOLD, ALTERNATIVE SPLICING, SIGNAL, 3 MUSCLE PROTEIN
423	1tit		353	437	5.1e-09	0.43	0.99		TITIN, I27; CHAIN: NULL;	IMMUNOGLOBULIN-LIKE DOMAIN CONNECTIN I27, TITIN IG REPEAT 27; MUSCLE PROTEIN, IMMUNOGLOBULIN-LIKE DOMAIN
423	1tit		357	435	9.9e-11	0.34	0.95		TITIN, I27; CHAIN: NULL;	IMMUNOGLOBULIN-LIKE DOMAIN CONNECTIN I27, TITIN IG REPEAT 27; MUSCLE PROTEIN, IMMUNOGLOBULIN-LIKE DOMAIN
423	1tit		644	722	6.6e-12	0.17	0.40		TITIN, I27; CHAIN: NULL;	IMMUNOGLOBULIN-LIKE DOMAIN CONNECTIN I27, TITIN IG REPEAT 27; MUSCLE PROTEIN, IMMUNOGLOBULIN-LIKE DOMAIN
423	1mm		353	429	6.8e-14	-0.02	0.52		MUSCLE PROTEIN TITIN MODULE M5 (CONNECTIN) 1TNM 3 (NMR, MINIMIZED AVERAGE STRUCTURE) 1TNM 4 1TNM 58	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
423	1tnm		667	722	6.6e-09	0.01	0.33		MUSCLE PROTEIN TITIN MODULE M5 (CONNECTIN) 1TNM 3 (NMR, MINIMIZED AVERAGE STRUCTURE) 1TNM 4 1TNM 58	
423	1vca	A	645	802	9.9e-11	0.15	0.42		HUMAN VASCULAR CELL ADHESION MOLECULE-1; 1VCA 4 CHAIN: A, B; 1VCA 5	CELL ADHESION PROTEIN VCAM-D1,2; 1VCA 6 IMMUNOGLOBULIN SUPERFAMILY, INTEGRIN-BINDING 1VCA 15
423	1wio	A	737	888	6.6e-13	0.20	0.86		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC LIPOPROTEIN, POLYMORPHISM
423	2dli	A	637	802	6.6e-12	0.12	-0.08		MHC CLASS I NK CELL RECEPTOR PRECURSOR; CHAIN: A;	IMMUNE SYSTEM P58 NATURAL KILLER CELL RECEPTOR; KIR, NATURAL KILLER RECEPTOR, INHIBITORY RECEPTOR, 2 IMMUNOGLOBULIN
423	2dli	A	741	888	2.6e-15	0.08	0.01		MHC CLASS I NK CELL RECEPTOR PRECURSOR; CHAIN: A;	IMMUNE SYSTEM P58 NATURAL KILLER CELL RECEPTOR; KIR, NATURAL KILLER RECEPTOR, INHIBITORY RECEPTOR, 2 IMMUNOGLOBULIN
423	2fcb	A	661	816	9.9e-16	0.10	0.09		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
423	2fcb	A	737	888	1.3e-15	0.04	0.01		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
423	2ncm		359	430	1.7e-08	-0.15	0.68		NEURAL CELL ADHESION MOLECULE; CHAIN: NULL;	CELL ADHESION NCAM DOMAIN 1; CELL ADHESION, GLYCOPROTEIN, HEPARIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										BINDING, GPI-ANCHOR, 2 NEURAL ADHESION MOLECULE, IMMUNOGLOBULIN FOLD, SIGNAL
423	7fab	L	733	863	1.7e-13	-0.02	0.05		IMMUNOGLOBULIN IMMUNOGLOBULIN FAB' NEW (LAMBDA LIGHT CHAIN) 7FAB 3	
423	8fab	A	733	863	6.8e-15	-0.11	0.12		IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	
424	1aut	L	113	206	3.3e-06			50.32	ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE), PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
424	1c2a	A	23	136	3.3e-17	-0.05	0.53		BOWMAN-BIRK TRYPSIN INHIBITOR; CHAIN: A	HYDROLASE INHIBITOR ALL-BETA STRUCTURE, HYDROLASE INHIBITOR
424	1en2	A	136	221	6.6e-07	-0.43	0.16		AGGLUTININ ISOLECTIN I/AGGLUTININ ISOLECTIN V/ CHAIN: A;	SUGAR BINDING PROTEIN UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPERANTIGEN, SACCHARIDE BINDING
424	1ext	A	88	250	1.6e-11			58.93	TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN, BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
424	1klo		22	185	6.6e-15			72.24	LAMININ; CHAIN: NULL;	GLYCOPROTEIN
424	1klo		30	226	6.6e-15	0.09	-0.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
424	4m2		120	185	1.7e-08			52.76	METALLOTHIONEIN METALLOTHIONEIN ISOFORM II 4MT2 3	GLYCOPROTEIN
424	9wga	A	11	185	1.7e-18			85.55	LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
425	1a06		116	394	8.5e-72			75.71	CALCIUM/CALMODULIN- DEPENDENT PROTEIN KINASE; CHAIN: NULL;	KINASE KINASE, SIGNAL TRANSDUCTION, CALCIUM/CALMODULIN
425	1a06		64	393	8.5e-72	-0.14	0.16		CALCIUM/CALMODULIN- DEPENDENT PROTEIN KINASE; CHAIN: NULL;	KINASE KINASE, SIGNAL TRANSDUCTION, CALCIUM/CALMODULIN
425	1a60		35	401	6.8e-40			73.63	PROTEIN KINASE CK2/ALPHA-SUBUNIT; CHAIN: NULL;	TRANSFERASE TRANSFERASE, SERINE/THREONINE- PROTEIN KINASE, CASEIN KINASE, 2 SER/THR KINASE
425	1apm	E	35	401	3.4e-75			78.39	TRANSFERASE(PHOSPHO TRANSFERASE) \$C-/AMP\$- DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (\$C/APK\$) 1APM 3 (CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (S139AS) COMPLEX WITH THE PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 1APM 6	
425	1apm	E	50	393	3.4e-75	0.33	0.92		TRANSFERASE(PHOSPHO TRANSFERASE) \$C-/AMP\$- DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (\$C/APK\$) 1APM 3	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									(CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 IAPM 4 REPLACED BY ALA (S139A) COMPLEX WITH THE PEPTIDE IAPM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 IAPM 6	
425	1aq1		66	395	1.7e-70	-0.04	0.89		CYCLIN-DEPENDENT PROTEIN KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; PROTEIN KINASE, CELL CYCLE, PHOSPHORYLATION, STAUROSPORINE, 2 CELL DIVISION, MITOSIS, INHIBITION
425	1aq1		66	401	1.7e-70			78.41	CYCLIN-DEPENDENT PROTEIN KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; PROTEIN KINASE, CELL CYCLE, PHOSPHORYLATION, STAUROSPORINE, 2 CELL DIVISION, MITOSIS, INHIBITION
425	1b6c	B	47	401	6.6e-31			80.91	FK506-BINDING PROTEIN; CHAIN: A, C, E, G; TGF-B SUPERFAMILY RECEPTOR TYPE I; CHAIN: B, D, F, H;	COMPLEX (ISOMERASE/PROTEIN KINASE) FKBP12; SERINE/THREONINE-PROTEIN KINASE RECEPTOR R4; COMPLEX (ISOMERASE/PROTEIN KINASE), RECEPTOR 2 SERINE/THREONINE KINASE
425	1bi8	A	123	393	6.8e-51			87.11	CYCLIN-DEPENDENT KINASE 6; CHAIN: A, C; CYCLIN-DEPENDENT KINASE INHIBITOR; CHAIN: B, D;	COMPLEX (KINASE/INHIBITOR) CDK6 P19INK4D; CYCLIN DEPENDENT KINASE, CYCLIN DEPENDENT KINASE INHIBITORY 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
425	1blx	A	62	401	3.4e-53			101.24	CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
425	1cki	A	62	398	1.6e-34			66.20	CASEIN KINASE I DELTA; 1CKI 6 CHAIN: A, B; 1CKI 7	PHOSPHOTRANSFERASE
425	1cmk	E	26	401	1.7e-76			82.38	PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT 1CMK 3 (E.C.2.7.1.37) 1CMK 4	PROTEIN KINASE 1CKI 18
425	1cmk	E	50	393	1.7e-76	0.19	0.95		PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT 1CMK 3 (E.C.2.7.1.37) 1CMK 4	
425	1csn		63	393	3.4e-15			62.13	CASEIN KINASE-I; 1CSN 4	PHOSPHOTRANSFERASE
425	1ctp	E	32	401	1.7e-76			90.28	TRANSFERASE(PHOSPHOTRANSFERASE) CAMP-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (CAPK) 1CTP 3 (CATALYTIC SUBUNIT) 1CTP 4	
425	1ctp	E	50	393	1.7e-76	0.18	0.94		TRANSFERASE(PHOSPHOTRANSFERASE) CAMP-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (CAPK) 1CTP 3	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									(CATALYTIC SUBUNIT) ICTP 4	
425	1f3m	C	49	393	5.1e-68	-0.05	0.06		SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: A, B; SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: C, D; FGF RECEPTOR 1; CHAIN: A, B;	TRANSFERASE KINASE DOMAIN, AUTOINHIBITORY FRAGMENT, HOMODIMER
425	1fgk	A	63	401	6.6e-31			84.95		PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2 PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
425	1fgk	B	54	401	3.4e-30			82.99	FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2 PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
425	1hcl		66	395	8.5e-75	0.09	0.92		HUMAN CYCLIN-DEPENDENT KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; TRANSFERASE, SERINE/THREONINE PROTEIN KINASE, ATP-BINDING, 2 CELL CYCLE, CELL DIVISION, MITOSIS, PHOSPHORYLATION
425	1hcl		66	401	8.5e-75			97.32	HUMAN CYCLIN-DEPENDENT KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; TRANSFERASE, SERINE/THREONINE PROTEIN KINASE, ATP-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
425	lian		50	401	6.8e-43			66.27	P38 MAP KINASE; CHAIN: NULL;	BINDING, 2 CELL CYCLE, CELL DIVISION, MITOSIS, PHOSPHORYLATION SERINE/THREONINE-PROTEIN KINASE CSBP, RK, P38; PROTEIN SER/THR-KINASE, SERINE/THREONINE-PROTEIN KINASE
425	lir3	A	54	393	9.9e-30			77.78	INSULIN RECEPTOR; CHAIN: A; PEPTIDE SUBSTRATE; CHAIN: B;	COMPLEX (TRANSFERASE/SUBSTRATE) TYROSINE KINASE, SIGNAL TRANSDUCTION, PHOSPHOTRANSFERASE, 2 COMPLEX (KINASE/PEPTIDE SUBSTRATE/ATP ANALOG), ENZYME, 3 COMPLEX (TRANSFERASE/SUBSTRATE)
425	lkoa		66	393	1.7e-56	0.19	0.94		TWITCHIN; CHAIN: NULL;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
425	lkob	A	43	400	1e-58			71.13	TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
425	lkob	A	59	393	1e-58	0.24	0.76		TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
425	lp38		50	401	3.4e-50			76.87	MAP KINASE P38; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE; TRANSFERASE, MAP KINASE, SERINE/THREONINE-PROTEIN KINASE, 2 P38
425	lphk		65	378	3.4e-67			75.55	PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE, GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE-PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
425	1phk		67	395	3.4e-67	0.19	0.93		PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE-PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING
425	1pme		60	401	1.7e-49			70.74	ERK2; CHAIN: NULL;	TRANSFERASE MAP KINASE, SERINE/THREONINE PROTEIN KINASE, TRANSFERASE
425	1tki	A	63	401	1.2e-46			57.34	TTTN; CHAIN: A, B;	SERINE KINASE SERINE KINASE, TTTN, MUSCLE, AUTOINHIBITION
425	3erk		55	401	3.4e-50			71.41	EXTRACELLULAR REGULATED KINASE 2; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE, MAP 2, ERK2; TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, MAP KINASE, 2 ERK2
426	1a7q	L	20	130	5.1e-29			61.78	MONOCLONAL ANTIBODY D1.3; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, VARIANT
426	1ivl	A	20	131	5.1e-28			62.21	IMMUNOGLOBULIN IMMUNOGLOBULIN VL DOMAIN (VARIABLE DOMAIN OF KAPPA LIGHT 1VL 3 CHAIN) OF DESIGNED ANTIBODY M29B 1VL 4	
426	1nfd	B	20	196	3.4e-51			88.42	N15 ALPHA-BETA T-CELL RECEPTOR; CHAIN: A, B, C, D; H57 FAB; CHAIN: E, F, G, H	COMPLEX (IMMUNORECEPTOR/IMMUNOGLOBULIN) COMPLEX (IMMUNORECEPTOR/IMMUNOGLOBULIN)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
426	2he		21	139	6.8e-31			61.07	IMMUNOGLOBULIN BENCE-JONES PROTEIN (LAMBDA, VARIABLE DOMAIN) 2RHE 4	
428	1g0u	D	69	295	3.4e-48	0.34	0.25		PROTEASOME COMPONENT Y7; CHAIN: A, O; PROTEASOME COMPONENT Y13; CHAIN: B, P; PROTEASOME COMPONENT PRE6; CHAIN: C, Q; PROTEASOME COMPONENT PUP2; CHAIN: D, R; PROTEASOME COMPONENT PRES; CHAIN: E, S; PROTEASOME COMPONENT C1; CHAIN: F, T; PROTEASOME COMPONENT C7-ALPHA; CHAIN: G, U; PROTEASOME COMPONENT PUP1; CHAIN: H, V; PROTEASOME COMPONENT PUP3; CHAIN: I, W; PROTEASOME COMPONENT C11; CHAIN: J, X; PROTEASOME COMPONENT PRE2; CHAIN: K, Y; PROTEASOME COMPONENT C5; CHAIN: L, Z; PROTEASOME COMPONENT PRE4; CHAIN: M, I;	HYDROLASE MACROPAIN SUBUNIT Y7, PROTEINASE YSCE SUBUNIT 7, MACROPAIN SUBUNIT Y13, PROTEINASE YSCE SUBUNIT 13, MACROPAIN SUBUNIT PRE6, PROTEINASE YSCE SUBUNIT MACROPAIN SUBUNIT PUP2, PROTEINASE YSCE SUBUNIT MACROPAIN SUBUNIT PRE5, PROTEINASE YSCE SUBUNIT MACROPAIN SUBUNIT C1, PROTEINASE YSCE SUBUNIT 1, MACROPAIN SUBUNIT C7- ALPHA, PROTEINASE YSCE MACROPAIN SUBUNIT PUP1, PROTEINASE YSCE SUBUNIT MACROPAIN SUBUNIT PUP3, MULTICATALYTIC MACROPAIN SUBUNIT C11, PROTEINASE YSCE SUBUNIT 11, MACROPAIN SUBUNIT PRE2, PROTEINASE YSCE SUBUNIT MULTICATALYTIC ENDOPEPTIDASE COMPLEX SUBUNIT C5; MACROPAIN SUBUNIT PRE4,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
428	1ryp	C	61	306	1.7e-47			76.38	PROTEASOME COMPONENT PRE3; CHAIN: N, 2;	PROTEINASE YSCE SUBUNIT MACROPAIN SUBUNIT PRE3, PROTEINASE YSCE SUBUNIT PROTEASOME, UBIQUITIN, DEGRADATION, PROTEASE, NTN-HYDROLASE
428	1ryp	C	62	292	1.7e-47	0.52	0.55		20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
428	1ryp	E	66	295	1.7e-48	0.31	0.58		20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
428	1ryp	E	67	300	1.7e-48			66.44	20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
428	1ryp	I	95	273	3.4e-42	0.56	1.00		20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	PROTEASE MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
428	1ryp	I	95	328	3.4e-42			117.29	20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
428	1ryp	L	95	295	8.5e-46	0.63	1.00		20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
428	1ryp	L	95	305	3.3e-56			201.96	20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
428	1ryp	L	95	309	3.3e-56	0.76	1.00		20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										PROCESSING, HYDROLASE, PROTEASE
430	1alh	A	471	120	1e-28	0.01	-0.01		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
430	1alh	A	498	573	1.2e-25	0.05	-0.14		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
430	1ard		498	526	1.7e-06	-0.20	0.64		TRANSCRIPTION REGULATION YEAST TRANSCRIPTION FACTOR ADR1 (RESIDUES 102 - 130) 1ARD 3 (AMINO TERMINAL ZINC FINGER DOMAIN) (NMR, 10 STRUCTURES) 1ARD 4 (ADR1B) 1ARD 5	
430	1bbo		471	520	1.7e-11	-0.41	0.24		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11 1BBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) 1BBO 4	
430	1bbo		474	520	1.3e-13	-0.47	0.65		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11 1BBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) 1BBO 4	
430	1buo	A	47	170	3.4e-19	0.40	0.89		PROMYELOCYTIC	GENE REGULATION POZ

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									LEUKEMIA ZINC FINGER PROTEIN PLZF; CHAIN: A;	DOMAIN; PROTEIN-PROTEIN INTERACTION DOMAIN, TRANSCRIPTIONAL 2 REPRESSOR, ZINC-FINGER PROTEIN, X-RAY CRYSTALLOGRAPHY, 3 PROTEIN STRUCTURE, PROMYELOCYTIC LEUKEMIA, GENE REGULATION
430	1mey	C	364	467	1.4e-45	0.11	-0.20		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
430	1mey	C	470	120	1e-47	-0.06	0.22		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
430	1mey	C	497	573	6.8e-42	0.07	-0.14		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
430	1mey	G	468	494	1.4e-11	-0.01	0.82		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
430	1mey	G	495	522	1.4e-13	0.02	0.75		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
430	1sp2		498	526	3.4e-09	-0.10	0.17		SPIF2; CHAIN: NULL;	ZINC FINGER TRANSCRIPTION FACTOR SP1; ZINC FINGER, TRANSCRIPTION ACTIVATION, SP1
430	1ubd	C	477	573	1e-30	0.08	-0.15		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
430	2adr		474	524	5.1e-17	-0.43	0.40		ADR1; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADR1, ZINC FINGER, NMR
430	7znf		498	526	1e-05	-0.62	0.16		ZINC FINGER DNA BINDING DOMAIN ZINC-FINGER (ZFY-SWAP) (NMR, 12 STRUCTURES) 7ZNF 3	
433	1dt6	A	55	497	0	0.77	1.00		CYTOCHROME P450 2C5; CHAIN: A;	OXIDOREDUCTASE PROGESTERONE 21-HYDROXYLASE, CYP1C5 P450 1. MEMBRANE PROTEIN, PROGESTERONE 21-HYDROXYLASE, BENZO(A) 2 PYRENE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
										HYDROXYLASE, ESTRADIOL 2-HYDROXYLASE, P450, CYP2C5
435	1alh	A	244	337	1e-23	-0.10	0.39		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
435	1alh	A	286	365	1.7e-26	0.06	0.74		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
435	1alh	A	544	628	5.1e-30	-0.10	0.75		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
435	1alh	A	572	655	6.8e-26	-0.24	0.59		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
435	1alh	A	830	894	1.7e-23	-0.24	0.72		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
435	1alh	A	842	923	1e-25	0.00	0.71		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									BINDING SITE; CHAIN: B, C;	PROTEIN
435	1alh	A	928	1030	3.4e-24	-0.17	0.04		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
435	1alh	A	978	1058	1e-26	0.04	0.36		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
435	1bbo		546	594	8.5e-11	-0.40	0.46		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11 IBBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) IBBO 4	
435	1bbo		548	596	6.6e-14	-0.42	0.45		DNA-BINDING PROTEIN HUMAN ENHANCER-BINDING PROTEIN MBP-1 MUTANT WITH CYS 11 IBBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) IBBO 4	
435	1mey	C	243	337	3.4e-42	-0.03	0.59		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
435	1mey	C	281	365	5.1e-45	-0.04	0.99		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER PROTEIN-DNA INTERACTION, PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
435	1mey	C	511	596	1.7e-40	-0.40	0.42		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
435	1mey	C	543	628	1.7e-48	-0.23	0.39		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
435	1mey	C	828	894	6.8e-35	-0.40	0.83		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
435	1mey	C	841	925	1.2e-43	-0.15	0.39		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
435	1mey	C	927	1030	5.1e-42	-0.19	0.03		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
435	1mey	C	977	1058	3.4e-45	0.03	0.86		DNA; CHAIN: A, B, D, E;	COMPLEX (ZINC

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
435	1mey	G	279	309	1e-10	-0.58	0.23		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
435	1mey	G	456	483	1.4e-08	0.27	-0.19		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
435	1mey	G	597	628	8.5e-11	-0.57	0.01		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
435	1mey	G	976	1002	1.2e-11	0.44	0.94		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
435	1uf3	A	544	628	3.4e-19	-0.16	0.45		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: B, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA, 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, TF

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
435	1tt6	A	512	657	1e-29	-0.24	0.09		TFIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
435	1ubd	C	246	365	5.1e-29	-0.34	0.80		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
435	1ubd	C	484	596	5.1e-26	-0.45	0.60		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
435	1ubd	C	515	628	6.8e-32	-0.39	0.36		YY1; CHAIN: C; ADENOVIRUS P5 ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
435	1ubd	C	551	655	6.8e-32	-0.23	0.36		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
435	1ubd	C	611	691	6.8e-12	0.06	-0.19		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
435	1ubd	C	837	921	1.7e-26	-0.21	0.70		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
435	2adr		546	598	3.4e-17	-0.25	0.88		ADRI; CHAIN: NULL;	TRANSCRIPTION REGULATION-TRANSCRIPTION REGULATION, ADRI, ZINC FINGER, NMR
435	2adr		572	630	3.4e-15	-0.34	0.24		ADRI; CHAIN: NULL;	TRANSCRIPTION REGULATION, ADRI, ZINC FINGER, NMR
435	2drp	A	545	596	9.9e-15	-0.05	0.19		COMPLEX(TRANSCRIPTIO	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
435	2drp	A	567	628	2.6e-16	-0.46	0.33		N REGULATION/DNA) TRAMTRACK PROTEIN (TWO ZINC-FINGER PEPTIDE) COMPLEXED WITH 2DRP 3 DNA 2DRP 4	
435	2drp	A	570	627	6.8e-08	-0.30	0.52		COMPLEX(TRANSCRIPTIO N REGULATION/DNA) TRAMTRACK PROTEIN (TWO ZINC-FINGER PEPTIDE) COMPLEXED WITH 2DRP 3 DNA 2DRP 4	
435	2gli	A	250	394	8.5e-25	-0.44	0.58		COMPLEX(TRANSCRIPTIO N REGULATION/DNA) TRAMTRACK PROTEIN (TWO ZINC-FINGER PEPTIDE) COMPLEXED WITH 2DRP 3 DNA 2DRP 4 ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE- FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA- BINDING PROTEIN/DNA)
435	2gli	A	516	657	1.5e-29	-0.26	0.65		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE- FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA- BINDING PROTEIN/DNA)
874	1f88	B	83	300	1.4e-11	-0.47	0.01		RHODOPSIN; CHAIN: A, B	SIGNALING PROTEIN PHOTORECEPTOR, G PROTEIN-COUPLED RECEPTOR, MEMBRANE PROTEIN, 2 RETINAL PROTEIN, VISUAL PIGMENT
437	1afc	A	38	174	6.8e-33			51.39	GROWTH FACTOR ACIDIC FIBROBLAST GROWTH FACTOR (AFGF) MUTANT WITH CYS 47 1AFC.3	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
437	1afc	A	45	173	6.8e-33	0.38	0.64		REPLACED BY ALA (C47A) COMPLEX WITH SUCROSE OCTASULFATE 1AFC 4	
437	1bar	B	29	174	1.7e-33			55.74	GROWTH FACTOR ACIDIC FIBROBLAST GROWTH FACTOR (AFGF) MUTANT WITH CYS 47 1AFC 3 REPLACED BY ALA (C47A) COMPLEX WITH SUCROSE OCTASULFATE 1AFC 4	
437	1bar	B	32	173	1.7e-33	0.56	0.76		GROWTH FACTOR ACIDIC FIBROBLAST GROWTH FACTOR (AFGF) MUTANT WITH CYS 47 1BAR 3 REPLACED BY ALA AND HIS 93 REPLACED BY GLY (C47A,H93G) 1BAR 4	
437	1bfg		38	164	3.4e-36			56.94	GROWTH FACTOR ACIDIC FIBROBLAST GROWTH FACTOR (AFGF) MUTANT WITH CYS 47 1BAR 3 REPLACED BY ALA AND HIS 93 REPLACED BY GLY (C47A,H93G) 1BAR 4	
437	1bfg		44	172	3.4e-36	0.61	0.46		GROWTH FACTOR BASIC FIBROBLAST GROWTH FACTOR MUTANT WITH CYS 69 REPLACED 1BFG 3 BY SER AND CYS 87 REPLACED BY SER (C69S,C87S) 1BFG 4	
437	1bla		11	164	1.7e-39			67.74	GROWTH FACTOR BASIC FIBROBLAST GROWTH FACTOR MUTANT WITH CYS 69 REPLACED 1BFG 3 BY SER AND CYS 87 REPLACED BY SER (C69S,C87S) 1BFG 4	GROWTH FACTOR FGF-2;

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQ FOLD score	Compound	PDB annotation
									GROWTH FACTOR; CHAIN: NULL;	GROWTH FACTOR
437	1bla		16	172	1.7e-39	0.11	0.03		BASIC FIBROBLAST GROWTH FACTOR; CHAIN: NULL;	GROWTH FACTOR FGF-2; GROWTH FACTOR
437	1qpk	A	39	172	3.4e-41	0.40	0.76		FIBROBLAST GROWTH FACTOR 7; CHAIN: A, B;	HORMONE/GROWTH FACTOR BETA-TREFOIL
437	1qpk	B	39	172	1.5e-37	0.29	0.25		FIBROBLAST GROWTH FACTOR 7; CHAIN: A, B;	HORMONE/GROWTH FACTOR BETA-TREFOIL
437	1qql	A	39	173	5.1e-37	0.62	0.75		FIBROBLAST GROWTH FACTOR 7/1 CHIMERA; CHAIN: A;	HORMONE/GROWTH FACTOR BETA-TREFOIL, HORMONE/GROWTH FACTOR
437	2afg	A	45	173	1.2e-33	0.49	0.66		ACIDIC FIBROBLAST GROWTH FACTOR; 2AFG 4 CHAIN: A, B, C, D; 2AFG 5	GROWTH FACTOR AFG; 2AFG 6

Table 6

SEQ.ID NO:	Position of Signal Peptide	Maximum score	Mean score
1	24	0.978	0.760
2	32	0.995	0.681
3	37	0.979	0.718
4	18	0.925	0.822
5	28	0.939	0.749
6	41	0.989	0.690
7	26	0.960	0.674
8	16	0.973	0.925
9	24	0.978	0.760
10	18	0.887	0.579
11	42	0.977	0.587
12	21	0.966	0.848
13	25	0.993	0.954
14	28	0.909	0.664
16	23	0.913	0.597
17	42	0.978	0.689
18	21	0.930	0.662
19	45	0.985	0.714
20	37	0.992	0.855
21	31	0.947	0.775
22	20	0.979	0.911
24	30	0.924	0.720
25	26	0.974	0.824
26	28	0.982	0.649
28	16	0.912	0.705
29	27	0.957	0.652
30	22	0.968	0.844
31	23	0.952	0.812
32	18	0.932	0.884
33	29	0.991	0.729
34	26	0.939	0.709
35	29	0.961	0.842
36	16	0.951	0.777
37	27	0.983	0.898
38	17	0.991	0.955
39	33	0.977	0.822
40	17	0.989	0.969
41	30	0.936	0.679
42	24	0.993	0.810
44	22	0.990	0.921
54	18	0.925	0.822
56	18	0.981	0.951
60	28	0.939	0.749
62	33	0.979	0.757
70	41	0.989	0.690
79	26	0.960	0.674
83	18	0.979	0.963
84	22	0.967	0.792
87	25	0.980	0.867
97	16	0.973	0.925
98	24	0.978	0.760
99	17	0.978	0.925

Table 6

SEQ.ID NO:	Position of Signal Peptide	Maximum score	Mean score
113	18	0.887	0.579
115	18	0.952	0.670
120	42	0.977	0.587
137	21	0.966	0.848
140	25	0.993	0.954
153	28	0.909	0.664
156	18	0.954	0.747
174	23	0.913	0.597
175	20	0.986	0.936
178	42	0.978	0.689
180	32	0.929	0.583
184	21	0.979	0.941
192	21	0.930	0.662
200	45	0.985	0.714
212	37	0.992	0.855
225	24	0.971	0.882
228	20	0.979	0.911
237	17	0.982	0.964
251	13	0.918	0.692
252	13	0.918	0.692
256	20	0.912	0.693
257	20	0.912	0.693
260	26	0.974	0.824
262	18	0.965	0.833
267	25	0.956	0.765
288	16	0.912	0.705
289	18	0.896	0.634
290	19	0.966	0.897
294	18	0.991	0.973
295	20	0.906	0.580
299	27	0.957	0.652
307	19	0.983	0.871
310	22	0.968	0.844
320	23	0.952	0.812
324	27	0.982	0.911
327	18	0.983	0.941
328	18	0.932	0.884
332	27	0.990	0.923
335	45	0.983	0.793
336	45	0.983	0.793
346	29	0.991	0.729
354	22	0.978	0.877
363	26	0.939	0.709
364	22	0.966	0.843
375	29	0.961	0.842
379	16	0.951	0.777
401	44	0.975	0.876
407	33	0.977	0.822
417	17	0.989	0.969
418	23	0.974	0.799
422	18	0.981	0.952
426	21	0.982	0.912

Table 6

SEQ.ID NO:	Position of Signal Peptide	Maximum score	Mean score
428	30	0.936	0.679
429	43	0.978	0.712
433	28	0.993	0.948
434	43	0.930	0.624
437	24	0.993	0.810
438	16	0.978	0.939

Table 7

SEQ ID NO:	Chromosomal location
3	2q11.2
4	20pter-p12.3
5	5q31
6	19p12
7	19p12
8	5
11	12p13-p12
12	p11.2-12.3
13	19p
14	6p12.1-21.1
15	19p13.1
17	16q12-q13
19	15
20	15
22	Xq13.1
23	12
25	11p15.5
26	20
27	22
28	12q23-24.1
29	20
30	13
31	12
33	15
36	4q28
37	14q24.3
38	10
39	20
41	17q12-q21
42	14
44	1q24.1-25.2
45	2
47	3q21-q25
48	9
49	14
50	6q14.1-15
51	19
52	11
53	20
54	16
55	14
56	3
57	19
58	7p15.1-p13
59	19
61	2
62	19
63	16
66	15
70	1p31.1-33
71	9
72	16

Table 7

SEQ ID NO:	Chromosomal location
74	5q31-q33
75	3p21.1-q13.13
76	2
77	2
78	21q22.1
79	Xp11.22-p11.21
80	2
81	19
82	20
83	19p13.3
84	19
85	3
86	8
87	1p13
88	16
89	18q21.1-q22
90	11q13.1-q13.3
91	18p11.23-p11.21
92	17
93	10
94	3
95	x
96	6q14.2-16.1
97	1q21.2-22
98	1q21.2-22
99	6
102	8q22-q23
103	10p11.2
104	17
105	17
106	2
107	1
108	16
109	17q21.3-q22
110	11q
111	3p21.1-q13.13
112	16
113	5
114	9
115	3p13-q26.1
116	5
117	7q31
118	14
119	14
120	19
121	19
122	6q27
123	14
124	1q21-q22
125	6
126	17q25
127	15

Table 7

SEQ ID NO:	Chromosomal location
129	14q31
130	1p36.1
131	11
132	20
133	20p11.23-p11.21
134	1p32
135	2q31
136	X
138	12p13
139	9
140	p34.1-34.3
141	19q12
142	15q26
143	22q11.21
144	17q12
145	4p16.3
146	22
147	16p11.2
148	18q12
150	4
151	7p12-q11.21
152	14
153	14q32.33
155	1p34
156	16p13.3
157	12p13.3
158	5
159	8
160	19
161	4
162	1
163	11q23
164	3
165	12q22
168	19
170	1
171	18q12
173	7
174	13
175	2p23.3-q32.3
176	16
178	10
179	1q21-q25
180	19p13.3
181	1
184	1p35.1-36.23
185	1
186	18
187	3p13-q26.1
188	3
189	17
190	6

Table 7

SEQ ID NO:	Chromosomal location
193	11p15.5
194	14q32
195	12
196	10q24
198	1p36.1
199	5q22
200	11
201	2q31
202	17
206	Xp11.23
207	9q34
208	19
209	20
210	11q23
211	16p12
212	19q13.1
213	7p15
214	15
215	1p36.21-36.33
216	11
217	22q11.2
218	15
219	19q13.4
222	19
223	1q25.2
226	1
227	1p36.11-36.23
228	1p36.3-p36.13
230	17
231	7q33-q34
232	3
233	9
234	10
235	17
236	4
237	19q13.4
238	4q25
239	2
240	7
241	12
243	6p21.3
244	3p13-q26.1
245	17
246	1p34.1
247	3q23
248	3p21.3
249	20
250	20
251	18q12-q21
252	18q12-q21
253	14
254	1p35.3-p35.1

Table 7

SEQ ID NO:	Chromosomal location
256	6q25-q26
257	6q25-q26
258	1q21-q23
259	16p13.2-16p13.11
260	14q21.1-q24.1
261	2p23.3-q32.3
262	12
263	19
264	4q28
265	2
266	2
267	1q21-q23
268	20p12.3-p13
269	4
270	6
271	2p23.3-q14.3
272	18q21
273	18q21
274	14q22
275	6p21.3
276	5
280	8
281	4q22-q24
282	2
283	7q22-q31.1
284	11
285	11q12.3
286	10
287	19
290	17
291	4q22
292	1p36.11-36.23
293	19
294	22
296	3
297	4p16
298	6
299	8q13
300	20
301	15
302	22q11.2-q22
303	15
304	6
306	6
307	9p24.2
308	2p23.3-q24.3
309	14
310	6
311	2
312	4
313	19pter-19p13.3
314	3

Table 7

SEQ ID NO:	Chromosomal location
316	11p12-14.2
317	19
318	17
319	17
320	5q14
323	4
324	3p
325	6p21.1-21.31
326	17p11.2
327	9
328	5q23
329	2
330	3
331	1p21.1-22.1
332	9
333	7
334	11q13
337	14
338	7q35-q36
339	13
340	6q11.1-22.33
341	11q12-q13.1
343	10
344	16
345	16
346	11q22
347	19
348	15q24-q26
350	Xp11.21-11.22
354	16
355	19
356	11
358	Xp11.23
359	4
360	8
362	4
363	11
364	11q13
365	7q31
366	22q13.31-13.32
367	5
370	19
371	7q31.1-7q31.33
372	2q37.3
373	3
374	16
375	19q13.4
376	18q12
377	18q12
379	8
380	11q13
381	6

Table 7

SEQ ID NO:	Chromosomal location
385	4q28
386	15
387	10
388	17
389	11p15.4
390	6p21.3
391	22q13
392	3
393	19
394	15
395	1
396	6p21.2-p21.3
397	15
399	7q31
400	14
402	Xq28
403	10
404	16
406	16
408	11
412	20q12-13.1
413	15
414	17
415	4
416	12q
419	21q22.1
420	16p11.2
422	6
424	21
426	14
428	14
429	1q22-q23
430	11q13
431	3
432	2
433	19q13.1
434	20q13.1
435	18q23
436	11q24
437	10
438	4q21-q25

Table 8

SEQ ID NO: of Full-length Nucleotide Sequence ¹	SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: in Priority Application USSN 09/774,528
52	52	54
53	53	55
54	54	56
55	55	57
56	56	58
57	57	59
58	58	60
59	59	61
60	60	62
61	61	63
62	62	64
63	63	65
64	64	66
65	65	67
66	66	68
67	67	69
68	68	70
69	69	71
70	70	72
71	71	73
72	72	74
73	73	75
74	74	76
75	75	77
76	76	78
77	77	79
78	78	80
79	79	81
80	80	82
81	81	83
82	82	84
83	83	85
84	84	86
85	85	87
86	86	88
87	87	89
88	88	90
89	89	91
90	90	92
91	91	93
92	92	94
93	93	95
94	94	96
95	95	97
96	96	98
97	97	99
98	98	100
99	99	101
100	100	102
101	101	103
102	102	104
103	103	105

Table 8

SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: in Priority Application USSN 09/774,528
104	104	106
105	105	107
106	106	108
107	107	109
108	108	110
109	109	111
110	110	112
111	111	113
112	112	114
113	113	115
114	114	116
115	115	117
116	116	118
117	117	119
118	118	120
119	119	121
120	120	122
121	121	123
122	122	124
123	123	125
124	124	126
125	125	127
126	126	128
127	127	129
128	128	130
129	129	131
130	130	132
131	131	133
132	132	134
133	133	135
134	134	136
135	135	137
136	136	138
137	137	139
138	138	140
139	139	141
140	140	142
141	141	143
142	142	144
143	143	145
144	144	146
145	145	147
146	146	148
147	147	149
148	148	150
149	149	151
150	150	152
151	151	153
152	152	154
153	153	155
154	154	156
155	155	157

Table 8

SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: in Priority Application USSN 09/774,528
156	156	158
157	157	159
158	158	160
159	159	161
160	160	162
161	161	163
162	162	164
163	163	165
164	164	166
165	165	167
166	166	168
167	167	169
168	168	170
169	169	171
170	170	172
171	171	173
172	172	174
173	173	175
174	174	176
175	175	177
176	176	178
177	177	179
178	178	180
179	179	181
180	180	182
181	181	183
182	182	184
183	183	185
184	184	186
185	185	187
186	186	188
187	187	189
188	188	190
189	189	191
190	190	192
191	191	193
192	192	194
193	193	195
194	194	196
195	195	197
196	196	198
197	197	199
198	198	200
199	199	201
200	200	202
201	201	203
202	202	204
203	203	205
204	204	206
205	205	207
206	206	208
207	207	209

Table 8

SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: in Priority Application USSN 09/774,528
208	208	210
209	209	211
210	210	212
211	211	213
212	212	214
213	213	215
214	214	216
215	215	217
216	216	218
217	217	219
218	218	220
219	219	221
220	220	222
221	221	223
222	222	224
223	223	225
224	224	226
225	225	227
226	226	228
227	227	229
228	228	230
229	229	231
230	230	232
231	231	233
232	232	234
233	233	235
234	234	236
235	235	237
236	236	238
237	237	239
238	238	240
239	239	241
240	240	242
241	241	243
242	242	244
243	243	245
244	244	246
245	245	247
246	246	248
247	247	249
248	248	250
249	249	251
250	250	252
251	251	253
252	252	254
253	253	255
254	254	256
255	255	257
256	256	258
257	257	259
258	258	260
259	259	261

Table 8

SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: in Priority Application USSN 09/774,528
260	260	262
261	261	263
262	262	264
263	263	265
264	264	266
265	265	267
266	266	268
267	267	269
268	268	270
269	269	271
270	270	272
271	271	273
272	272	274
273	273	275
274	274	276
275	275	277
276	276	278
277	277	279
278	278	280
279	279	281
280	280	282
281	281	283
282	282	284
283	283	285
284	284	286
285	285	287
286	286	288
287	287	289
288	288	290
289	289	291
290	290	292
291	291	293
292	292	294
293	293	295
294	294	296
295	295	297
296	296	298
297	297	299
298	298	300
299	299	301
300	300	302
301	301	303
302	302	304
303	303	305
304	304	306
305	305	307
306	306	308
307	307	309
308	308	310
309	309	311
310	310	312
311	311	313

Table 8

SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: in Priority Application USSN 09/774,528
312	312	314
313	313	315
314	314	316
315	315	317
316	316	318
317	317	319
318	318	320
319	319	321
320	320	322
321	321	323
322	322	324
323	323	325
324	324	326
325	325	327
326	326	328
327	327	329
328	328	330
329	329	331
330	330	332
331	331	333
332	332	334
333	333	335
334	334	336
335	335	337
336	336	338
337	337	339
338	338	340
339	339	341
340	340	342
341	341	343
342	342	344
343	343	345
344	344	346
345	345	347
346	346	348
347	347	349
348	348	350
349	349	351
350	350	352
351	351	353
352	352	354
353	353	355
354	354	356
355	355	357
356	356	358
357	357	360
358	358	361
359	359	362
360	360	363
361	361	364
362	362	365
363	363	366

Table 8

SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: in Priority Application USSN 09/774,528
364	364	367
365	365	368
366	366	369
367	367	370
368	368	371
369	369	372
370	370	373
371	371	374
372	372	375
373	373	376
374	374	377
375	375	378
376	376	379
377	377	380
378	378	381
379	379	382
380	380	383
381	381	384
382	382	385
383	383	386
384	384	387
385	385	388
386	386	389
387	387	390
388	388	391
389	389	392
390	390	393
391	391	394
392	392	395
393	393	396
394	394	397
395	395	398
396	396	399
397	397	400
398	398	401
399	399	402
400	400	403
401	401	404
402	402	405
403	403	406
404	404	407
405	405	408
406	406	409
407	407	410
408	408	411
409	409	412
410	410	413
411	411	414
412	412	415
413	413	416
414	414	417
415	415	418

Table 8

SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: in Priority Application USSN 09/774,528
416	416	419
417	417	420
418	418	421
419	419	422
420	420	423
421	421	424
422	422	425
423	423	426
424	424	427
425	425	428
426	426	429
427	427	430
428	428	431
429	429	432
430	430	433
431	431	434
432	432	435
433	433	436
434	434	437
435	435	438
436	436	439
437	437	440
438	438	441

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-438, a mature protein coding portion of SEQ ID NO: 1-438, an active domain coding portion of SEQ ID NO: 1-438, and complementary sequences thereof.
2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
6. A vector comprising the polynucleotide of claim 1.
7. An expression vector comprising the polynucleotide of claim 1.
8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:

- (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and
 - (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO: 1-438.
- 11. A composition comprising the polypeptide of claim 10 and a carrier.
- 12. An antibody directed against the polypeptide of claim 10.
- 13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
 - a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and
 - b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.
- 14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
 - a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;
 - b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
 - c) detecting said product and thereby the polynucleotide of claim 1 in the sample.
- 15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.
- 16. A method for detecting the polypeptide of claim 10 in a sample, comprising:

a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and

b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.

17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and

b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

19. A method of producing the polypeptide of claim 10, comprising,

a) culturing a host cell comprising a polynucleotide sequence selected from SEQ ID NO: 1-438, a mature protein coding portion of SEQ ID NO: 1-438, an active domain coding portion of SEQ ID NO: 1-438, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-438, under conditions sufficient to express the polypeptide in said cell; and

b) isolating the polypeptide from the cell culture or cells of step (a).

20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of any one of the polypeptides encoded by SEQ ID NO: 1-438, the mature protein portion thereof, or the active domain thereof.
21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.
22. A collection of polynucleotides, wherein the collection comprising the sequence information of at least one of SEQ ID NO: 1-438.
23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.
24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.
25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.
26. The collection of claim 22, wherein the collection is provided in a computer-readable format.
27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.
28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.